NISTIR 6524

Information Models for Design Tolerancing: From Conceptual to the Detail Design

R. Sudarsan[1]
U. Roy[2]
Y. Narahari[3]
R.D. Sriram[4]
K.W. Lyons[5]
N. Pramanik[6]

[1]NIST (Visiting)
Gaithersburg, MD 20899 and
The George Washington University
Washington, DC 20052
[2 6]Mechanical, Aerospace, Manufacturing
Engineering Department
Syracuse University
Syracuse, NY 13244
[3]Department of Computer Science
and Automation
Indian Institute of Science
Bangalore, 560 012, India and
NIST
Gaithersburg, MD 20899 (Visiting)
[4 5]U.S. DEPARTMENT OF COMMERCE
Design and Process Group
Manufacturing Systems Integration Division
National Institute of Standards
and Technology
Gaithersburg, MD 20899

NIST
National Institute of Standards and Technology
Technology Administrotion, U.S. Department of Commerce

NISTIR 6524

Information Models for Design Tolerancing: From Conceptual to the Detail Design

R. Sudarsan[1]
U. Roy[2]
Y. Narahari[3]
R.D. Sriram[4]
K.W. Lyons[5]
N. Pramanik[6]

[1]NIST (Visiting)
Gaithersburg, MD 20899 and
The George Washington University
Washington, DC 20052
[2 6]Mechanical, Aerospace, Manufacturing
Engineering Department
Syracuse University
Syracuse, NY 13244
[3]Department of Computer Science
and Automation
Indian Institute of Science
Bangalore, 560 012, India and
NIST
Gaithersburg, MD 20899 (Visiting)
[4 5]U.S. DEPARTMENT OF COMMERCE
Design and Process Group
Manufacturing Systems Integration Division
National Institute of Standards
and Technology
Gaithersburg, MD 20899

May 2000

U.S. DEPARTMENT OF COMMERCE
William M. Daley, Secretary

TECHNOLOGY ADMINISTRATION
Dr. Cheryl L. Shavers, Under Secretary
of Commerce for Technology

NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
Raymond G. Kammer, Director

Information Models for Design Tolerancing: From Conceptual to the Detail Design

R. Sudarsan [1] U. Roy [2] Y.Narahari [3] R.D. Sriram [4] K.W. Lyons [5]
N. Pramanik [6]

[1] The George Washington University, Washington, DC 20052 and the National Institute of Standards and Technology, Gaithersburg, MD 20899 (Visiting). Email: sudarsan@cme.nist.gov

[2] Mechanical, Aerospace, Manufacturing Engineering Department, Syracuse University, Syracuse, NY 13244, Email: uroy@ecs.syr.edu

[3] Department of Computer Science and Automation, Indian Institute of Science, Bangalore, 560 012, India and the National Institute of Standards and Technology, Gaithersburg, MD 20899 (Visiting). Email: hari@csa.iisc.ernet.in

[4] Engineering Design Technologies Group, Manufacturing Systems Integration Division, National Institute of Standards and Technology, Gaithersburg, MD 20899. Email: sriram@cme.nist.gov

[5] Engineering Design Technologies Group, Manufacturing Systems Integration Division, National Institute of Standards and Technology, Gaithersburg, MD 20899. Email: klyons@darpa.mil

[6] Mechanical, Aerospace, Manufacturing Engineering Department, Syracuse University, Syracuse, NY 13244, Email: pramanik@ecs.syr.edu

Contents

List of Figures

List of Tables

SYMBOLS AND NOTATION

C_p	Process capability index without mean shift
C_{pk}	Process capability index with mean shift
C_m	Taguchi Process capability index
C_{pi}, C_{pki}	Process capability indices for individual parts or processes
$i = 1, \ldots, n$	Index for individual dimensions
$l(Y)$	Loss function (Loss when Y deviates from τ)
L	Expected quadratic loss, $E[l(Y)]$
LSL	Lower specification limit of Y
$Pr\{X\}$	Probability of X
T	Tolerance of Y
T_i	Tolerance of X_i
USL	Upper specification limit of Y
X_i	Dimension or characteristic of an individual part
Y	Assembly response characteristic (such as assembly gap or any function)
$Y_1, \ldots, Y_m$	Individual assembly response characteristics
$Y = f(X_1, \ldots, X_n)$	Assembly response function
τ	Target value of Y
μ, μ_Y	Expected value of Y
μ_i	Expected value of X_i
σ, σ_Y	Standard deviation of Y
σ_i	Standard deviation of X_i
CAD	Computer Aided Design
CAM	Computer Aided Manufacturing
CAPP	Computer Aided Process Planning
CBR	Case Based Reasoning
CDF	Cumulative Density Function
DFA	Design for Assembly
DFC	Datum Flow Chain
DFT	Design for Tolerancing
dof	degrees of freedom
DRF	Datum Reference Frame
ERP	Enterprise Resource Planning
FAB	Function Assembly Behavior
GPS	Geometrical Product Specification
HPD	Holistic Probability Design

OO	Object Oriented
OpenADE	Open Assembly Design Environment
PCD	Process Capability Data
PDM	Product Data Management
PFM	Part Function Model
QFD	Quality Function Deployment
STEP	STandard for the Exchange of Product model data
UML	Unified Modeling Language
VR	Virtual Reality

Webpages

ANSI	American National Standards Institute	http://www.ansi.org
ASME	American Society of Mechanical Engineers	http://www.asme.org
EDT	Engineering Design Technologies Group	http://www.mel.nist.gov/ msid/groups/edt/index.html
ISO	International Organization for Standardization	http://www.iso.ch
ISO TC 213	ISO Technical Committee 213	http://www.ds.dk/isotc213
NIST	National Institute of Standards and Technology	http://www.nist.gov
OMG	Object Management Group	http://www.omg.org
RaDEO	Rapid Design Exploration and Optimization	http://www.darpa.mil/ DSO/rd/Manufact/rapid.html
SIMA	System Integration for Manufacturing Applications	http://www.mel.nist.gov/ msid/sima/sima.htm
SOLIS	STEP On-Line Information Service	http://www.nist.gov/sc4

Information Models for Design Tolerancing: From Conceptual to the Detail Design

R. Sudarsan, U. Roy, Y. Narahari, Ram D. Sriram, K.W.Lyons, N. Pramanik

Abstract:

Tolerance design is the process of deriving a description of geometric tolerance specifications for a product from a given set of desired properties of the product. Existing approaches to tolerance analysis and synthesis entail detailed knowledge of geometry of assemblies and are mostly applicable during advanced stages of design, leading to a less than optimal design. During the design process of assemblies, both assembly structure and associated tolerance information evolve continuously and significant gains can be achieved by effectively using this information to influence the design of an assembly. Any pro-active approach to the assembly or tolerance analysis in the early design stages will involve making decisions with incomplete information models. In order to carry out early tolerance synthesis and analysis in the conceptual product design stage, we need to devise techniques for representing function-behavior-assembly models that will allow analysis and synthesis of tolerances, even with the incomplete data set.

A 'function' (what the system is for) is associated with the transformation of an input physical entity into an output physical entity by the system. The problem or customer's need, initially described by functional requirements on an assembly and associated constraints on the functional requirements defines the concept of an assembly. This specification of functional requirements and constraints define a functional model for the assembly. Many researchers have studied functional representation (function based taxonomy and ontology), function to form mapping, and behavior representation (behavior means how the system/product works). In a recent paper, [68], we presented a strong need for comprehensive function-assembly-behavior (FAB) integrated model.

In this report, we discuss extension of the ideas presented in [68] and explain the integration of function, assembly, and behavior representation into a comprehensive information model (FAB models). To do this, we need to develop appropriate assembly models and tolerance models that would enable the designer to incrementally understand the build-up or propagation of tolerances (i.e., constraints) and optimize the layout, features, or assembly realizations. This will ensure ease of tolerance delivery. In an earlier paper, [53], a multi-level approach called Design for Tolerance [DFT] process was proposed which enables tolerancing to be addressed at successive stages of design in an incremental fashion We also address the effective use of the FAB and DFT model for design tolerancing, starting from conceptual stage of the design and continuously evolving throughout the entire design process to the final detailed design. These models can eventually lead to tolerance and assembly standards.

1 Introduction

Tolerancing is a critical issue in the design of electro-mechanical assemblies. Tolerancing includes both tolerance synthesis and tolerance analysis. There are several research thrusts that have been reported in individual areas of assembly modeling, tolerance analysis, and tolerance synthesis. But an integrated and complete information model that supports both assembly activities (i.e., assembly plan generation) and tolerancing synthesis and analysis process uniformly throughout the entire product design process does not yet exist. As outlined in [53], an ideal assembly model for tolerancing should: 1) be closely coupled with the design process; 2) be mutable through successive design stages; and 3) be capable of representing assembly and tolerance information at any level of abstraction.

The other important attributes that have been identified for an appropriate assembly model are as follows: 1) capturing design intent; 2) embedding different views (relational view, location logic, etc.) in the assembly information model; 3) enabling all assembly information including tolerancing to be captured in the model; and 4) representing effective, integrated assembly information models throughout the design process.

Tolerancing decisions can profoundly impact the quality and cost of electro-mechanical assemblies. Existing approaches to tolerance analysis and synthesis in design entail detailed knowledge of geometry of assemblies and are mostly applicable during advanced design stages, leading to a less than optimal design process. During the design process of assemblies, both assembly structure and associated tolerance information evolve continuously. Therefore, significant gains can be achieved by effectively using this information to influence the design of an assembly. Motivated by this, we identified and explored two goals for future research that we believe can enhance the scope of tolerancing for the entire design process [53, 54]. The first goal addresses advancing tolerance related decisions to the earliest possible design stages. This issue raises the need for effective representation of tolerancing information during different design stages and for effective assembly modeling. The second goal addresses appropriate, synergistic use of available methods and best practices for tolerance analysis and synthesis, at successive design stages. Pursuit of these goals lead us to the definition of a multi-level approach that enables tolerancing to be addressed at successive design stages in an incremental fashion. The resulting design process called the design for tolerance [53] process integrates three important domains: 1) design activities at successive design stages; 2) assembly models that evolve continuously through the design process; and 3) methods and best practices for tolerance analysis and synthesis.

Product functionality is another important factor to be considered. It is clear that any intelligent decision during the design of product modeling cannot be made without the knowledge of product functions. The information model for assembly must include functional and behavioral characteristics of its component parts.

Setting tolerances for parts of an assembly has always been a critical design decision. The design of tolerances for component parts and assembly should be cost effective and adequate

to ensure the required performance specifications for an assembly. An inappropriate choice of tolerances can result in low quality products, extensive or difficult manufacturing steps, or both.

For complex designs, tradition, trial and error, or intuition frequently determine tolerances. A common method employed by designers is to select the dimensions of a part that are considered important, and then specify tightest tolerances that manufacturing processes can achieve. This unnecessarily overburdens the manufacturing facilities without ensuring optimality of the design. We need to design tolerances that help in the rational choice based on considerations of cost, sensitivity, and performance specifications. However, we should remember that current assembly analysis tools or tolerance analysis/synthesis tools have always been utilized in the detailed design phase of the product development process.

In this report our main aim is to develop an integrated comprehensive function, assembly (artifact) and behavior (FAB)[1] model. To do this, we need to develop appropriate assembly models and tolerance models that would enable the designer to incrementally understand the build-up or propagation of tolerances and optimize the layout, features, or assembly realizations. This will ensure ease of tolerance delivery. Any pro-active approach to assembly or tolerance analysis in the early design stages will involve decision making with incomplete information models. In order to carry out early tolerance synthesis and assembly analysis in the conceptual design stage, we need to devise a means of representation of FAB data models that will allow analysis, even with the incomplete data set.

The report is organized as follows. In order to build the subject matter and explain the FAB and Design For Tolerancing (DFT) models we first explain the basic concepts and methods of tolerancing analysis and synthesis and describe the current and evolving standards for statistical tolerancing in Section 2. The proposed information model for FAB is discussed in Section 3. In Section 4, we present the integration of FAB and DFT models. In Section 5 we discuss about the current and evolving standards and the contribution from the proposed models (FAB and DFT). Suggestions for further research is presented in Section 6. Finally we present some conclusions in Section 7.

A brief tutorial on process capability indices is given in Appendix A. The defect rate of 3.4 parts per million (ppm) and the employed assumptions and convention is explained in Appendix B.

2 Methods of Tolerance Analysis and Synthesis

Tolerance design is a two-stage process consisting of tolerance synthesis and analysis. Tolerance synthesis is the process of allocating tolerances among the geometric parameters (dimen-

[1]We have chosen the name FAB model to draw the readers attention to the fact that the product design activity starts with the functional requirements, assembly (either a skeletal or list of parts), followed by a quick check on meeting the functional requirements by doing a behavioral simulation

sions) of the part/component such that the cost incurred due to tolerances is minimized and at the same time the functional and assembly requirements are satisfied. Tolerance analysis on the other hand, is concerned with the aggregate behavior of the individual tolerances.

2.1 Methods for Tolerance Analysis

Tolerance analysis can be either *worst-case* or *statistical*. Worst-case tolerance analysis-also called deterministic or high-low tolerance analysis-considers the worst possible combinations of individual tolerances and examines the assemblability of the parts, so as to achieve 100% interchangeability of parts in an assembly. This results in unnecessarily tight part tolerances and hence high production costs. Statistical tolerancing is a more practical and economical way of looking at tolerances and works on setting the tolerances so as to assure a desired yield. Here, the designer abandons the notion of 100% interchangeability and accepts some small percent of non-conformance.

In tolerance analysis, we seek to ascertain the validity of the specified part tolerances with regard to assemblability and manufacturability. Tolerance analysis can be carried out by determining the tolerance zones [65, 66] belonging to the specified tolerance types. It can be either statistical tolerance analysis (where statistical methods are used together with accompanying probability distribution) or worst case analysis (i.e. the study of extreme cases). The FAB data model as discussed in Section 3.3.1 (page: 38) provides all the required information for tolerance analysis as the data model can be viewed as both a functional and a structural hierarchy. In order to carry out tolerance analysis in the early design stages when geometry is ill defined (i.e., incomplete) the functional hierarchy becomes very helpful.

The tolerance analysis of the toleranced parts in assembly shows how a structure of the part-assembly satisfies the constraints of design and manufacture, and how the position of parts or their features are constrained by the tolerance specification. Tolerance models based on well-defined variational models are needed to help analyze the assembly configuration uncertainties. Mathematical formalization of assembly conditions is needed to show how contacts between tolerance parts are modeled. The FAB data model provides explicit descriptions of functions of the various components and their spatial geometric relationships, including the tolerancing information.

Statistical tolerance analysis uses a relationship of the form:

$$Y = f(X_1, \ldots, X_n)$$

where Y is the response (a measurable characteristic such as assembly gap) of the assembly and $X_1, \ldots, X_n$ are the values of some characteristics (such as dimensions) of the individual parts or sub-assemblies making up the assembly. We call f the assembly response function (ARF). The relationship can exist in any form for which it is possible to compute a value for Y given values of $X_1, \ldots, X_n$. It could be any of the following: an explicit or an implicit

analytic expression; a complex engineering calculations; result from experiments and/or simulations. The input variables $X_1, \ldots, X_n$ are continuous random variables. In general, they could be mutually dependent. The function f is a deterministic relationship; Y is a continuous random variable. The general problem of tolerance analysis is to compute the probability distribution of Y given the distributions of $X_1, \ldots, X_n$. However, more often we are usually interested in computing the first few moments (mean, standard deviation, skewness, and kurtosis), given the distributions or first few moments of the input variables. Once the moments of Y are determined, one can compute a tolerance range for Y that would envelope a given fraction of the assembly yield.

There are a variety of methods and techniques available for the above computational problem. Essentially, the methods can be categorized into four classes [18]:

1. Stack Tolerancing or Linear Propagation (Root Sum of Squares)
2. Non-linear propagation (Extended Taylor series)
3. Numerical integration (Quadrature technique)
4. Monte Carlo simulation

2.1.1 Linear Propagation

This is also called stack tolerancing and uses the well-known root sum of squares (RSS) formula. The assembly response function here is of the form:

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

where $a_0, a_1, \ldots, a_n$ are constants and $X_1, \ldots, X_n$ are assumed to be mutually independent. Many dimensional and gap-related measures fall into this category. Because of the linear relationship and mutual independence, the mean and variance of Y are given by:

$$\begin{aligned} \mu_Y &= a_0 + a_1\mu_1 + a_2\mu_2 + \ldots + a_n\mu_n \\ \sigma_Y^2 &= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots + a_n^2\sigma_n^2 \end{aligned}$$

where μ_i is the mean and σ_i, the standard deviation of X_i, $i = 1, \ldots, n$. The nomenclature RSS arises because of the above formula for standard deviation. If the individual distributions are normal, then Y is also normally distributed. Even if the individual distributions are not normal, Y can safely be treated as normal, by invoking the central limit theorem.

If the linear relation for Y above is only approximately true, then one can expand $f(X_1, \ldots, X_n)$ as a Taylor series and drop all but the constant and linear terms. This is often-used device in statistical tolerancing to handle approximately linear relationships. In such a case,

$$a_i = \frac{\partial f}{\partial x_i} \quad \text{evaluated at} \quad x_i = \mu_i, \quad i = 1, \ldots, n,$$

and all of the constant terms are gathered into a_0. The computation of the above partial derivatives could be of two types. In the first case, the function f is known and the partial derivatives are known to exist. In the second case, the functional relationship is either too intractable or not even available in analytic form. In such a case, numerical estimates have to be obtained for the partial derivatives [18].

The linear case is the simplest and the most efficient among all tolerance analysis approaches. It is very appealing for synthesis methods that use analysis in an iterative way.

2.1.2 Non-linear Propagation (Extended Taylor Series)

If the assembly response function Y is highly non-linear, application of the RSS method could lead to serious errors. An extended Taylor series approximation for the relationship f can possibly be employed provided f is available in an analytic form. Usually, the expansion is considered up to the sixth order. The expansion is possible only when all the appropriate partial derivatives exist. The main computational issue involves computing the partial derivatives of f. Mathematically tractable formulae for the first four moments of Y are available [18] and are ideally suited for tolerance analysis and synthesis. These formulae need only the first four moments of the distributions of $X_1, \ldots, X_n$. Most often, the partial derivatives are computed using analytic methods. If closed form solutions do not exist then we need to use numerical iterations techniques as described in the next section.

2.1.3 Numerical Integration

If the function f is not available in analytic form and Y can only be computed through numerical calculations or engineering methods or simulations, numerical methods have to be used. Quadrature methods are generally used here. Here, we assume that for any function $h(X_1, \ldots, X_n)$ (different from f) of mutually independent random variables $X_1, \ldots, X_n$ with probability density functions $w_{X_i}(x_i)$, the expected value of h is given by the integral

$$\int_{-\infty}^{+\infty} \cdots \int h(x_1, \ldots, x_n) \prod_{i=1}^{n} (w_{X_i}(x_i) dx_i)$$

The above expression can be approximated by a quadrature expression [18] that involves evaluations of h at $2n^2 + 1$ prescribed values. These evaluations involve only the first four moments of $X_1, \ldots, X_n$. Given an assembly response function f, a corresponding function h as above can be defined and simple moment transfer relations can be used to compute the first four moments of f. The quadrature technique adapts well to statistical tolerancing problems since it can handle the iteration inherent in a tolerancing problem efficiently. An improved integration technique is provided in [55, 58].

2.1.4 Monte Carlo Simulation

The appeal of Monte Carlo simulation lies in its applicability under very general settings and the unlimited precision that can be achieved. In particular, Monte Carlo simulation can be used in all situations in which the above three techniques (stack tolerancing, extended Taylor series, and numerical integration) can be used and can yield more precise estimates. For this reason, the Monte Carlo technique is easily the most popular tool used in tolerancing problems. The caveat, however, is the large computational time: for situations where the above three techniques are adequate and have acceptable precision, the Monte Carlo technique is much more expensive in terms of computational time.

Monte Carlo simulation proceeds as follows (Figure 1). Pseudo random number generators are used to generate a sample set of numbers $x_1, \ldots, x_n$, belonging to the random variables $X_1, \ldots, X_n$, respectively. The value of Y, say $y_1 = f(x_1, \ldots, x_n)$, corresponding to this sample is computed. This procedure is replicated a large number of times, say N times. This would yield a random sample, $\{y_1, \ldots, y_N\}$, for Y. Standard statistical estimation methods are then used to analyze the distribution of Y. The precision of this statistical analysis increases as proportional to $\sqrt{N}$ and therefore unlimited precision can be achieved through large number of replications. Special techniques are available for significantly enhancing the precision of the Monte Carlo method for a given N. These include: weighted sampling, reuse of samples, and use of approximation functions [18].

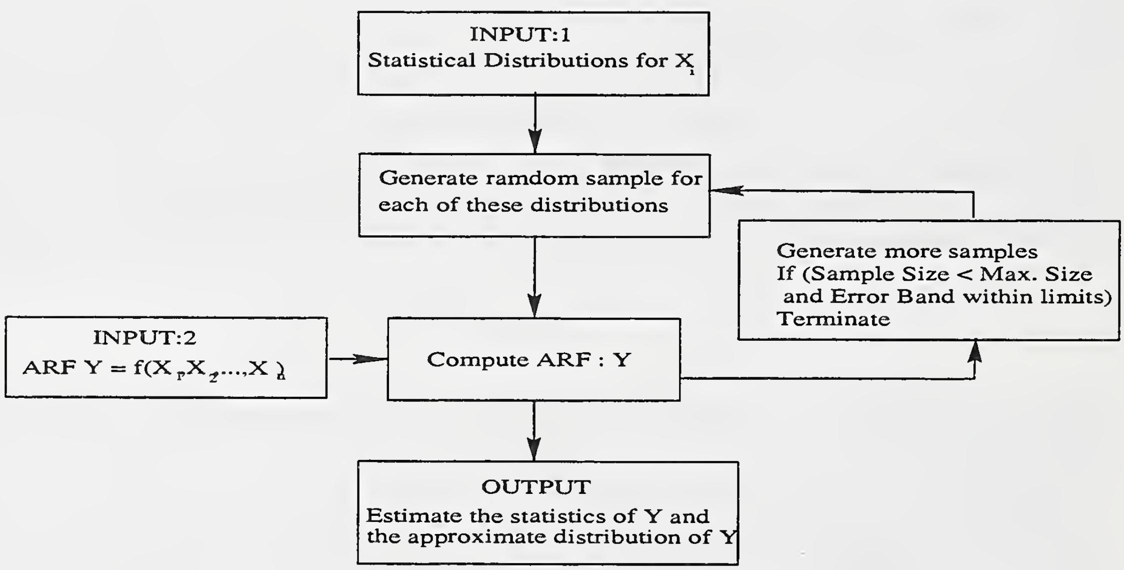


Figure 1: Flow chart of Monte-Carlo simulation method

2.2 Methods for Tolerance Synthesis

In the context of electro-mechanical assembly design, tolerance synthesis usually refers to the allocation of specified assembly tolerances among the constituent parts and sub-assemblies, so as to ensure a specified yield or minimize a proper cost function. More generally, if $Y = f(X_1, \ldots, X_n)$ is an assembly response function, then the synthesis problem involves finding the best nominals and tolerances for $X_1, \ldots, X_n$, given nominal and tolerance specifications for Y. Synthesis is naturally an optimization problem; one can formulate an objective function that captures yield requirements or production cost requirements and pose an optimization problem by including tolerance related constraints.

There are several views and variants of the synthesis problem, depending on the objective function and the constraints. One view is to minimize the total manufacturing cost by choosing both the individual part tolerances and the manufacturing processes for making the individual parts. This requires cost versus tolerance relationships for each individual dimension. Another view is to find *robust* nominals for individual dimensions, i.e., nominal values at which the effect of variations on the assembly response function is minimum. This issue is addressed by Taguchi's robust design methodology and Park's Holistic Probabilistic Design (HPD). Also, depending on the nature of the objective function and the constraints, the synthesis problem can be deterministic or stochastic.

To formulate the synthesis problem meaningfully, a certain amount of preprocessing is often required. For example, one needs to first determine the tolerance limits on the assembly response function, Y. An important preprocessing step is *sensitivity analysis*, which determines which assembly parameter variations have significant effects on the assembly response function. This reveals the set of parameters or individual dimensions to emphasize in the synthesis procedure.

2.2.1 Iterative Methods Based on Analysis

A simple and realistic mechanism for tolerance synthesis is to employ a trial and error technique for postulating tolerances for individual parts and sub-assemblies. Next, perform a statistical tolerance analysis required to ascertain whether this postulated set of tolerances fulfills the desired criteria. If the chosen set is unsatisfactory, some of the tolerances are changed and the analysis redone; this step is repeated until a satisfactory set of part tolerances is obtained. Typically, at the end of each iteration, we obtain a probability of assembly, a probability of conformance, an expected yield, or a more detailed cost. This technique effects trade-offs during each iteration and is appealing as it uses the findings of the current iteration to drive the next iteration. During early iterations, approximate cost figures and less accurate estimates can be used. These can be replaced by more accurate figures as the iterations start producing good solutions.

The methods discussed for statistical tolerance analysis, namely stack tolerancing, ex-

tended Taylor series, quadrature methods, and Monte Carlo, are all suited for the iterative approach. Evans [18] has discussed the merits and issues concerning the use of these methods for the iterative methodology for synthesis.

2.2.2 Multistage Tolerance Synthesis

In order to synthesize tolerance in the conceptual design stage, Roy and Bharadwaj [63] suggested a multi stage procedure. Figure 2 shows the conceptual schema for tolerance synthesis. Given the design function requirements, manufacturing processing information and assembly plan, the schema helps assign both dimensional and geometric tolerances (along with required datum reference planes) to be part of an assembly. The tolerance synthesis schema starts with collecting the following information (refer to Phase 1, Figure 2) from the FAB data model.

1. Geometry description: the schema requires geometry description at the following levels:
    - assembly - position and orientation information for each component within the assembly
    - part-spatial location of form features in the part and their inter-relationships and
    - feature-feature geometry.
2. Part function specifications: Functional and behavioral attributes of each part in the assembly.
3. Material and surface finish specifications:
    - Material and surface characteristics should be either retrieved from the database or supplied manually by the user.
4. Assembly graph:
    - The procedure for assembling different parts in the assembly (without considering the effect of tolerances) should be retrieved from the assembly data model. The assembly graph denotes the assembly sequence of the parts in order to form an assembly. It is a directed acyclic graph whose nodes are represented by "assembly states" and whose edges denote the "assembly process".

In the second phase, the part function model (PFM) and the process model for each part of the assembly are generated. A part function model describes the spatial and design relationships that exist on the mating faces of a part in terms of certain kinematics, force degrees of freedom (dofs), presence/absence of motions, and transmission of forces along

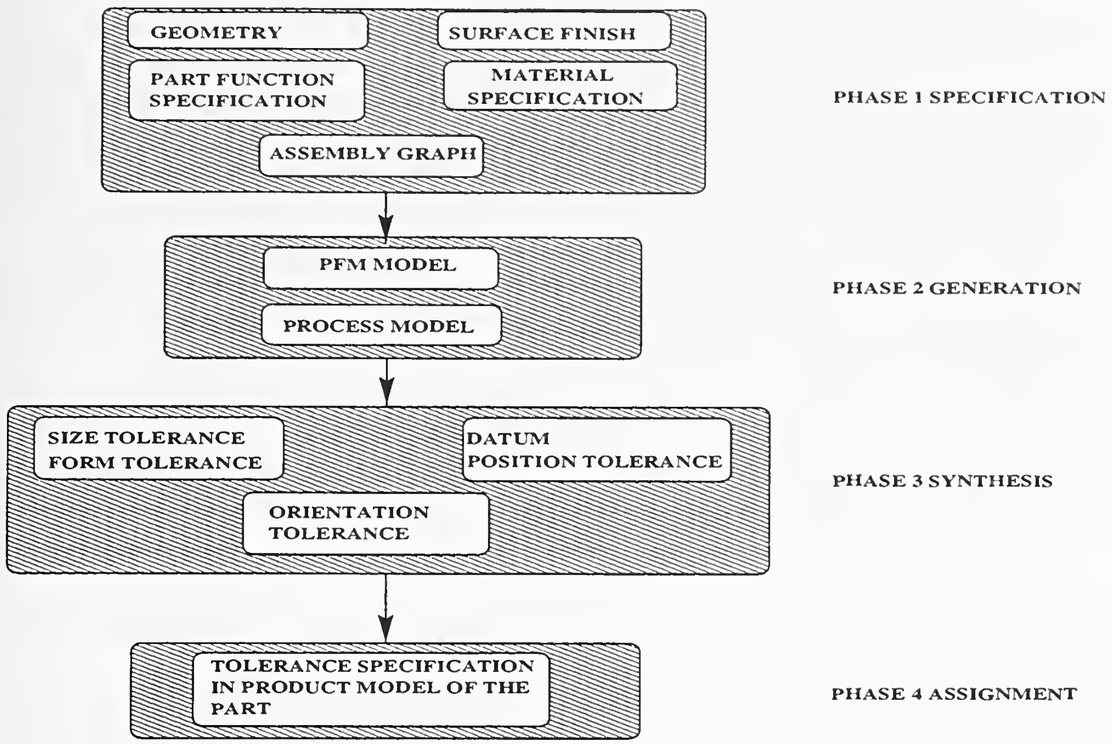


Figure 2: Tolerance Synthesis Scheme

the particular axes of a surface [7]. The process model represents the process plan for manufacturing the part without considering the effect of tolerance [4].

The third phase of the schema is the synthesis stage. Different types of tolerances are synthesized for each part of the assembly. It consists of two major tasks: (i) transformation of the PFM model into functional tolerance limits, and (ii) constraining the functional tolerance limits with respect to different manufacturability and assemblability constraints. The first task can be achieved by developing appropriate application domain-specific PFM-to-Functional-Tolerance-Limit maps (refer to [63, 64] for a detailed discussion). The second task can be achieved by developing optimization problems which contain both the functional tolerance limits and the different constraints. In the fourth phase, dimensional and geometric tolerances (along with the datum specifications) are finetuned with respect to the design functions and manufacturing constraints.

The FAB data model provides the required information needed for the above four stage tolerance synthesis process.

2.2.3 Optimization Methods

Since tolerance synthesis can be posed as an optimization problem, mathematical programming techniques such as linear programming, non-linear programming, and, integer programming are relevant. There have been several efforts in this direction [13, 43, 46]. Also, heuristic techniques for optimization such as simulated annealing, genetic algorithms, Lagrangian relaxation, and Tabu search have been used by researchers [25, 38, 45].

The following integer programming formulation given by Kusiak and Feng [43], provides a flavor of a typical synthesis problem in the optimization framework:
Minimize

$$\sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}x_{ij}$$

subject to:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} t_{ij}x_{ij} \leq T$$

$$\sum_{j=1}^{n} x_{ij} = 1 \quad \forall i$$

$$x_{ij} = 0, 1 \quad \forall i, j$$

where the index i denotes one of n dimensions involved in the assembly; index j denotes one of m manufacturing processes that can be used for producing a dimension; c_{ij} is the manufacturing cost of process j when used for producing dimension i; t_{ij} is the 3σ normal variation of process j when used to produce dimension i; T is the tolerance stack-up limit for the assembly; and x_{ij} is a binary decision variable that takes a value 1 if process j is selected for producing dimension i and 0 otherwise. Note that the objective is to minimize the total direct manufacturing cost, by choosing the appropriate tolerances and the right mix of manufacturing processes. In the above formulation, a linear relationship has been assumed between part tolerances and also worst case tolerancing has been used. Hence, it is a deterministic tolerance synthesis problem. In HPD, the optimization is done considering nominals and variance together.

2.2.4 Design of Experiments

Here, the assembly response function (or in general, a well-defined cost function), is computed for various discretized values of the random variables $X_1, \ldots, X_n$, (dimensions with tolerances) according to the design of experiments theory. The factors used in the experiment include not only the individual values of $X_1, \ldots, X_n$, but also factors that capture tolerance related constraints. Full factorial or fractional factorial designs can be used depending on the number of factors and levels of the factors. Prior sensitivity analysis can help in choosing the levels for the factors. The setting that leads to a minimum cost and also satisfies the

tolerance constraints can be chosen as the solution. One can go a step further and fit a statistical model that describes the cost function in terms of all the factors and use this model to arrive at an optimal solution for the problem.

Taguchi methods, which are described in the next section, use design of experiments in a novel way to find robust nominals.

2.3 Best Practices

In the last decade, many companies have established comprehensive programs in total quality management. These efforts include those of Motorola, Xerox, IBM, AT&T Bell Laboratories, and several others that have initiated formal, corporate programs for improved tolerance specification, monitoring, and control. In this section, we outline the tolerancing best practices developed at Motorola and Xerox which are closely related to our approach. We also provide an overview of Taguchi's robust design methodology, which has emerged as a best practice in many companies.

2.3.1 Motorola Six Sigma Program

Six sigma quality is the benchmark of excellence for product and process quality, popularized by Motorola [22, 24]. It provides a quantitative and statistical notion of quality useful in understanding, measuring, and reducing variation. A product is said to be of six sigma quality if there are no more than 3.4 non-conformities per million opportunities (3.4 ppm) at the part and process-step level, in the presence of typical sources of variation. The six sigma quality concept recognizes that variations are inevitable due to insufficient design margin, inadequate process control, imperfect parts, imperfect materials, fluctuations in environmental conditions, operator variations, etc.

Tolerance analysis and synthesis in the six sigma program are based on the following criteria:

1. *six sigma characterization of products and processes* where the process capability indices C_p and C_{pk} are used as the vehicles to characterize the product-process quality;

2. *simple, intuitive extensions to the RSS method* which enables tolerance analysis and synthesis in the presence of shifts and drifts of the process mean; and

3. *a well-defined, systematic program for design for quality* taking into account both the product perspective and the process perspective.

We provide a brief outline of these issues; for a detailed discussion refer [22–24]. Before discussing the tolerance analysis and synthesis using Motorola six-sigma methodology, we discuss the Motorola six-sigma quality program.

Definition of Motorola Six-Sigma Quality

According to the Motorola six-sigma program [22], a product is said to be of six-sigma quality (0.9999966 confidence) if there are *no* more than 3.4 non-conformities per million of opportunities, at the product and process level, in the presence of typical sources of variation (shifts and drifts). The six-sigma quality represents a 2941 (10000/3.4) fold improvement in quality compared to the 99% quality.

The six-sigma design uses the capability indices C_p and C_{pk} as quality metrics (refer Appendix A, page 67 for more details). The six-sigma design has a defect rate of 3.4 parts per million with $C_p \geq 2$ and $C_{pk} \geq 1.5$ under the following assumptions that the process (or a typical product characteristic) (1) is normally distributed and (2) has a mean shift of 1.5σ units from the target value. The explanation of the defect rate 3.4 ppm with aid of an example is given in Appendix B.

2.3.2 Tolerance Analysis and Synthesis in the Motorola Approach

The Motorola program assumes a linear model for Y, of the form

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

If there is no mean shift, then the standard RSS formulae are applicable:

$$\begin{aligned} \mu_Y &= a_0 + a_1\mu_1 + a_2\mu_2 + \ldots + a_n\mu_n \\ \sigma_Y^2 &= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots + a_n^2\sigma_n^2 \end{aligned}$$

Recall that σ_i, for $i = 1, \ldots, n$, can also be written as:

$$\sigma_i = \frac{T_i}{3C_{pi}}$$

where T_i is the tolerance range of the ith part and C_{pi} is the C_p value for the ith part ($i = 1, \ldots, n$). In the presence of a mean shift, the standard RSS formula cannot be used for computing the standard deviation of Y. Two alternative approaches are recommended by the Motorola program. The first is the Dynamic RSS (DRSS) where the C_{pk} values, $C_{pk1}, \ldots, C_{pkn}$, of the individual processes corresponding to dimensions $X_1, \ldots, X_n$ and the tolerances, $T_1, \ldots, T_n$, of the individual parts are used in the following way to compute the variance of Y:

$$\sigma_Y^2 = a_1^2\left(\frac{T_1}{3C_{pk1}}\right)^2 + \ldots + a_n^2\left(\frac{T_n}{3C_{pkn}}\right)^2$$

Note that the standard deviations σ_i are amplified by an amount equal to $\frac{C_{pi}}{C_{pki}}$, for $i = 1, \ldots, n$. Thus the DRSS method emulates random behavior in the process mean by inflating

the process standard deviation, but has little impact on the overall mean. The second alternative method called Static RSS (SRSS) achieves the sustained mean-shifts condition by applying a correction factor to the individual nominals. For details, see [24].

Tolerance analysis is carried out by using RSS, DRSS, and SRSS, as appropriate. Tolerance synthesis uses tolerance analysis in an iterative way. Each iteration will evaluate the resulting probability of non-conformance and the C_p and C_{pk} values. The goal of the synthesis procedure is to obtain a probability of non-conformance of at most 3.4 ppm, which is guaranteed by $C_p \geq 2$ and $C_{pk} \geq 1.5$. The synthesis can assume several forms:

1. finding optimal values for nominal dimensions;
2. finding optimal values for tolerances; and
3. establishing a variance pool that can be allocated to individual processes so as to obtain the desired assembly yield.

The Motorola six sigma approach uses the normal distribution for all its probability and tolerancing computations. While this can be listed as a limitation, it takes very little away from the intrinsic novelty and applicability of the approach. The ideas it has germinated essentially hold in all situations; only the probability computations need to be redone under non-normal situations and the quantitative measures need appropriate reinterpretation.

2.3.3 Holistic Probabilistic Design (HPD)

The HPD methodology [56–58] is one of several quality programs at the Xerox Corporation. The program is based on relating service dissatisfiers and customer tolerances to a set of critical parameters (parameters that are critical to the product's function). The tolerances of the critical parameters are related to piece part variabilities through multiple *flow of variance chains*. Tolerance analysis and synthesis are carried out through the chains to yield the desired quality. Since the objective is to maximize the amount of manufacturing and usage variability the product can tolerate, with negligible impact on the targeted level of performance, the program is also called *design for latitude*. The methodology is implemented using a complete suite of tools for stochastic variability analysis. These include the following:

1. A stochastic modeling based technique for computing the distribution of a function of a random variables almost exactly;
2. Contribution analysis that provides a reliable guidance on factors that have a significant effect on the assembly response function; and
3. An operating window optimization method that helps choose the densities of certain input random variables for which the allowable range of operation is maximized.

Tolerance analysis is based on a stochastic technique that uses a failure rate prediction model. Let $Y = f(X_1, \ldots, X_n)$ be the assembly response function, as usual. If failure is defined by $f(X_1, \ldots, X_n) > Y_0$, then the probability of failure is given by:

$$Pr\{f(X_1, \ldots, X_n) > Y_0\} = \int \underset{\Omega}{\ldots} \int \; w(X_1, \ldots, X_n) dX_1 \ldots dX_n$$

where Ω is the n-dimensional failure region and $w(X_1, \ldots, X_n)$ is the joint density of the n random variables. From this, it is easy to see how to compute the distribution function of Y. Assuming mutual independence of $X_1, \ldots, X_n$, the HPD tool uses an efficient numerical technique to evaluate such multiple integrals. This enables us to compute the distribution accurately. This computation is versatile since it can handle any type of distribution and any type of relationship; it has excellent computational performance if the number of random variables is less than 10. The above computation enables variability analysis, hence tolerance analysis. It also provides a sound basis for iterative tolerance synthesis. An attractive feature of this technique is its applicability to both geometric and non-geometric type of situations. For example, the random variables $X_1, \ldots, X_n$ need not be dimensions; they could be physical quantities such as force, pressure, temperature, and speed.

The contribution analysis embedded within HPD is a powerful feature of HPD. It provides a sound basis for determining the input variables that have a pronounced effect on the assembly response function. Also, it takes into account the nature of the input distributions and accounts for cross-term effects. This has several advantages over other existing techniques for contribution analysis.

Another feature of HPD is the stochastic operating window optimization. This feature enables us to maximize the allowable range of operation by intelligently selecting the densities of input random variables using tools provided by HPD.

The HPD tool consists of two major modules: HPD-VA and HPD-OPT. The module HPD-VA is a stochastic analyzer that includes variability analysis and tolerance analysis. HPD-OPT has a wide variety of deterministic, stochastic, and statistical optimization routines. HPD-OPT finds the most robust set of nominals and tolerances.

2.3.4 Taguchi Methods

Taguchi methods, also known as robust design methods, [36, 37, 59], are technical methods for quality and cost control at the product and process design stages. According to Taguchi, the cost of a product is the loss incurred to the society before the product is shipped to the customer, whereas quality is the loss imparted to the society after the product has been shipped to the customer. Such losses include the following: loss due to harmful side effects; loss due to variations in the product's performance characteristics; and all losses that can be traced to the poor performance of the product. Taguchi methods emphasize reducing

the sensitivity of engineering designs to various sources of variation. The methods are cost-effective as it minimizes the influence of variation sources rather than control them.

Let Y be a *performance characteristic*; as before, Y is a continuous random variable. Taguchi considers Y as a function of *design parameters* or *design factors*, $\Theta = (\theta_1, \ldots, \theta_k)$, and *noise parameters* or *noise factors*, $W = (w_1, \ldots, w_t)$. Thus, $Y = f(\Theta, W)$. Design factors are input variables whose nominal settings and that have a pronounced influence on Y can be chosen by a designer. Design factors are of two types: *signal factors*, which affect only the mean of Y, and *control factors*, which affect both mean and variance of Y. Noise factors are input variables that cause Y to deviate from its target value. Noise factors include deviations of the actual values of design factors from the nominal settings. Taguchi considers two types of matrices: the *design matrix* (inner array) and the *noise matrix* (outer array). The design matrix has k columns, each column corresponding to a particular design factor. Each row of this matrix represents a specific combination of design parameter settings. The number of rows depends on the number of combinations of design parameter settings being investigated. Similarly, the noise matrix has t columns, each column corresponding to a particular noise factor. Each row of this matrix represents a specific combination of noise factor settings. The number of rows depends on the number of combinations of noise factor settings sought to be investigated.

Let τ be the *target value* (ideal value) of Y; μ, its mean; and σ, its standard deviation. The target value need not be the midpoint of a tolerance interval. Variations of Y about the target value τ cause losses to the customers. Let $l(Y)$ be the loss due to the deviation of Y from τ. Taguchi suggests a quadratic form for the loss function:

$$l(Y) = k(Y - \tau)^2$$

where k is a constant that can be computed from a known value of loss at any designated value of Y. The loss function is a random variable and the expected quadratic loss, $E[l(Y)]$, is given by

$$L = E[l(Y)] = k(\sigma^2 + (\mu - \tau)^2)$$

Thus the expected quadratic loss is the the sum of the variance of Y and the square of bias (bias indicates how far away from the target value the process mean is behaving).

Minimization of the expected quadratic loss is the primary objective of Taguchi methods. This is accomplished by maximizing a signal-to-noise ratio (also called as a performance statistic). Taguchi's use of signal-to-noise ratios rather than directly employing the expected quadratic loss is motivated by considerations such as ease of statistical estimation, more direct coupling to design factors, and improved additivity of control factor effects. For a detailed exposition refer [37, 59]. A signal-to-noise ratio is a statistical estimate of the effect of noise factors on Y for a particular setting of design parameters. Numerous performance statistics have been defined by Taguchi (more than 60).

The following are the main steps in the Taguchi method.

1. Identify appropriate loss function or signal-to-noise ratio, initial and competing settings of design factors, and important noise factors and their ranges.

2. Construct the *design matrix* and the *noise matrix*. The design matrix is chosen based on the theory of design of experiments or is chosen from Taguchi's collection of orthogonal arrays [37]. The noise matrix is usually chosen from Taguchi's collection of orthogonal arrays.

3. Conduct a *parameter design experiment*. This involves N_d *runs*, where N_d is the number of rows of the design matrix. Each run corresponds to a particular row and involves N_n *replications*, where N_n is the number of rows of the noise matrix. For each run,i, $(i = 1, \ldots, N_d)$, a corresponding signal-to-noise ratio, $[Z(\Theta)]_i$ is computed.

4. Use $[Z(\Theta)]_1, \ldots, [Z(\Theta)]_{N_d}$, to predict a *statistical model* for the signal to noise ratio. Use the predicted statistical model to obtain optimal or best design parameter settings:

$$\Theta^* = (\theta_1^*, \ldots, \theta_k^*)$$

5. Conduct a verification experiment to confirm that Θ^* indeed minimizes the expected loss. Otherwise, iterate.

Taguchi methods make several assumptions. Examples of these include: absence of interaction effects among the factors; additivity of control factors; separability of signal factors and control factors; and use of signal-to-noise ratios instead of direct measures [37, 59]. However, the methodology embodies sound engineering considerations and intuition for obtaining robust designs, which explains its widespread use. From a tolerancing viewpoint, Taguchi methods provide a valuable tool for synthesizing robust nominals. Also, the methods can be applied potentially during early assembly design stages. Furthermore, the methods enable economic considerations to be incorporated into tolerance analysis and synthesis, and provide an approach that is complementary in many ways to other best practices such as the Motorola six sigma program and the Xerox HPD methodology.

2.4 Summary

In summary, in this section we discussed the various tolerance analysis and synthesis techniques, tools and industry best practices. In the next section we will discuss about the current and evolving statistical tolerancing standards syntax and semantics.

Although statistical tolerancing has been practiced in industry for a long time, standards are not yet available. ISO [34, 72, 73] is investigating how to standardize statistical tolerancing. The scope of ISO TC 213 [34] is standardization in the field of geometrical product specifications (GPS) i.e. macro- and micro-geometry specifications covering dimensional and

geometrical tolerancing, surface properties and the related verification principles, measuring equipment and calibration requirements including the uncertainty of dimensional and geometrical measurements. This standardization includes the basic layout and explanation of drawing indications (symbols).

2.5 Statistical Tolerancing

Both classical and statistical tolerancing are currently practiced in the industry. Classical tolerancing is popularly known as the worst-case tolerancing. Given an actual feature, it defines a test to decide whether the feature is acceptable. Statistical tolerancing, on the other hand, deals with a population of features. It defines a test that decides whether a given population of actual features is acceptable.

Let us consider the classic example discussed by Evans [17]. In that example a stack of n nominally identical disks is analyzed. For illustrative purpose, consider $n = 10$ and the height of the stack as the response (assembly response function $Y = X_1 + X_2 + \cdots + X_n$ [53]). Let us assume that the stack height must be 1.25 ± 0.01 ($USL_y = 1.26, LSL = 1.24$). The worst case tolerancing of the disks would yield 0.125 ± 0.001 for the components (individual disks).

The problem with statistically tolerancing the components of an assembly (individual disks) deals with setting limits on the allowed excursion of the components' mean and range (using the standard deviation). Let us assume that $X_i \sim N(\mu_i, \sigma_i)$ with the assumption that $\sigma_i = \sigma_c\ \forall i$. If σ_i^2 is the variance of the component i, the variance of assembly is given by

$$\sigma_y^2 = Var(\text{assembly}) = \sum \sigma_i^2 = n\sigma_c^2 \quad \text{if } \sigma_i = \sigma_c;\ \forall i \tag{1}$$

Component level to assembly Level

Traditionally, the half width of the specified tolerance is set to three times the standard deviation of the assembly ($C_p = 1$) in determining the standard deviation of each component. In this example, let $\mu_i = 0.125$ and $\sigma_i = \frac{0.001}{3}$ so that $\mu_y = \sum \mu_i = 1.25$ and $\sigma_y = \sqrt{10}\left(\frac{0.001}{3}\right)$. That is, given $X_i \sim N(0.125, \frac{0.001}{3})$ then the assembly variable $Y \sim N(1.25, \sqrt{10}\left(\frac{0.001}{3}\right))$. We can now compute the following probability:

$$Pr\{Y \in [1.24, 1.26]\} \approx Pr\{1.25 - 10\sigma_y \leq Y \leq 1.25 + 10\sigma_y\} \approx 10^{-20}$$

Assembly level tolerancing to component level

Now let us do the statistical analysis from the assembly level to component level. Set $\pm 3\sigma$ limits for the stack height as 0.01. Since $\sigma_y^2 = Var(\text{assembly}) = \sum \sigma_i^2 = n\sigma_c^2$, the process would be designed so that:

$$\sigma_c = \frac{0.01}{3\sqrt{10}}.$$

The tolerance on individual disks could be $3\sigma_c$ that is in terms of high/low tolerancing 0.125 ± 0.003. Thus approached from the viewpoint of statistical tolerancing, the tolerances on the individual disks can be relaxed by a factor of three.

There are various reasons for doing statistical tolerancing: (1) worst-case tolerancing leads to unnecessary tight tolerances, costing more money and time; (2) rate of miniaturization exceeds the rate of improvement in the manufacturing process capabilities; and (3) it is the missing link in Shewhart's quality control process [75].

2.5.1 Statistical Tolerancing - Syntax and Semantics

The boundary of an actual part (model of a manufactured part) can be partitioned into many actual surface features. Since these surfaces do not have perfect form, we numerically fit a perfect-form feature.

ISO TC 213 defines several types of features. For example a truncated portion of a cylindrical surface on a nominal geometrical model of a part is called *nominal integral feature* (see Figure 3) . Its axis is named the *nominal derived feature* to highlight the fact that the axis is derived from a cylindrical surface that is integral to the part boundary. The nominal geometrical model is only a concept in the mind of a designer. An actual realization of the part will have a surface that only approximately corresponds to the cylindrical feature, this is called the *real integral feature.* The real integral feature has an infinite number of points and for the purpose of measurements only a finite subset of these are used. These finite number of points sampled on the real integral surface form what is called the *extracted integral feature.* Then a perfect form surface, such as a full cylinder, is fitted to the sampled points to form what is termed the *associated integral feature.* In ISO/TC213, the term association is used synonymously with fitting. Finally, *associated derived feature* is the axis or other appropriate entity that is derived from the fitted surface.

Deriving a Distribution

As mentioned already we fit a perfect-form feature to an actual feature. As shown in Figure 4, two types of fits are popular: Gaussian and Chebychev fits.

A Gaussian fit minimizes the squares of deviations of the sensed points from the substitute element and the solution is unique. The Chebychev method, on the other hand, uses a min-max criteria. It minimizes the maximum deviation of the sensed points from the substitute element. The choice of fit may be dictated by the function of the part. For example, if assemblability is the major functional criteria, then Chebychev fit is preferred. The planned ISO 10360 [27] uses Gaussian fit because it gives a unique solution.

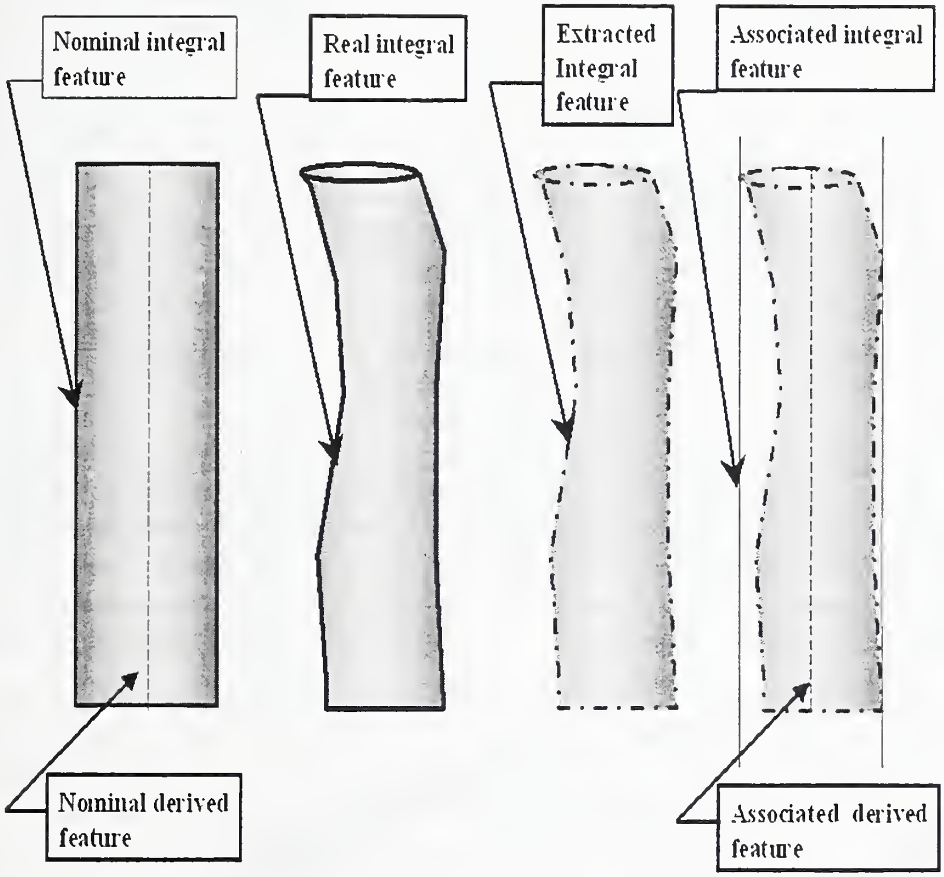


Figure 3: Various types of features defined by ISO TC 213 ([71])

Acceptable Distributions

In the previous section we discussed the process of mapping a population of actual features to the distribution of one or more random variables of interest. In this section we discuss the

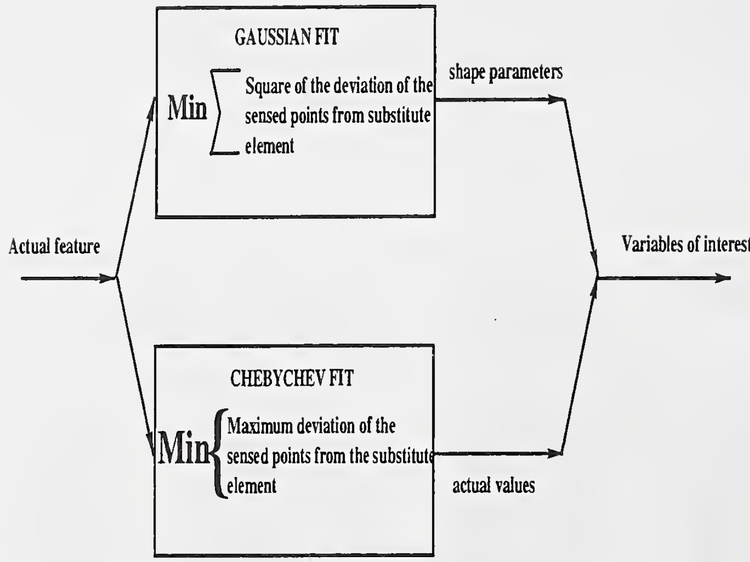


Figure 4: Extracting parameters or actual values by fitting ([72, 73])

acceptable class of population of actual features (parts), by defining which statistical distributions are acceptable. In [72, 73], three ways of specifying acceptable class of distributions are proposed by defining statistical tolerance zones. We briefly discuss them here. For more details, we refer the interested reader to [72, 73].

2.5.2 Statistical Tolerance Zones

The statistical tolerance zones are of two types: (1) parametric statistics, and (2) nonparametric statistics. In parametric statistics, the population parameter zones are defined in the space of parameters of the population. Currently, only the first two (central) moments of the distribution are considered in the population parameter space. In the nonparametric statistics, the distribution function zones are defined using the whole distribution. Conceptually, nonparametric methods are more general than parametric methods.

Parametric Statistics

Using the process capability indices explained in Appendix A (page: 67), a process is statistically controlled by specifying lower bounds for C_p and C_{pk}, and an upper bound for the centering index ($k \equiv C_c$, as defined in [72]). Essentially we need to specify three number P, K and F, so that

$$C_p \geq P \qquad \text{(lowerbound)} \tag{2}$$

$$
\begin{aligned}
C_{pk} &\geq K \qquad \text{(lowerbound)} \qquad (3)\\
C_c &\leq F \qquad \text{(upperbound)} \qquad (4)
\end{aligned}
$$

The condition 4 implies $C_{pk} \geq (1-F)C_p$. The following zones fall under this type:

1. $C_p - C_{pk}$ zone,
2. $\mu - \sigma$ zone, and,
3. C_{pm} zone.

Zone in $C_{pk} - C_p$ Plane

Each value of the triplet (P, K, F) corresponds to a tolerance zone in the $C_{pk} - C_p$ plane. We are interested in the first quadrant of $C_{pk} - C_p$ zone where both C_{pk} and C_p are positive and $0 \leq F \leq 1$. The tolerance zone is the intersection of the four half-planes

$$
\begin{aligned}
C_p &\geq P\\
C_{pk} &\geq K\\
C_{pk} - C_p &\leq 0\\
C_{pk} &\geq (1-F)C_p \qquad (5)
\end{aligned}
$$

Figure 5 shows the tolerance zone for the general case of (P, K, F) and the Motorola-six sigma case $(P = 2, K = 1.5, F = 0.25)$. In case we do not have lower bound for C_{pk}, we can use the condition given by the equation (5) above. In such cases, the shape of the zone is similar to the six-sigma case.

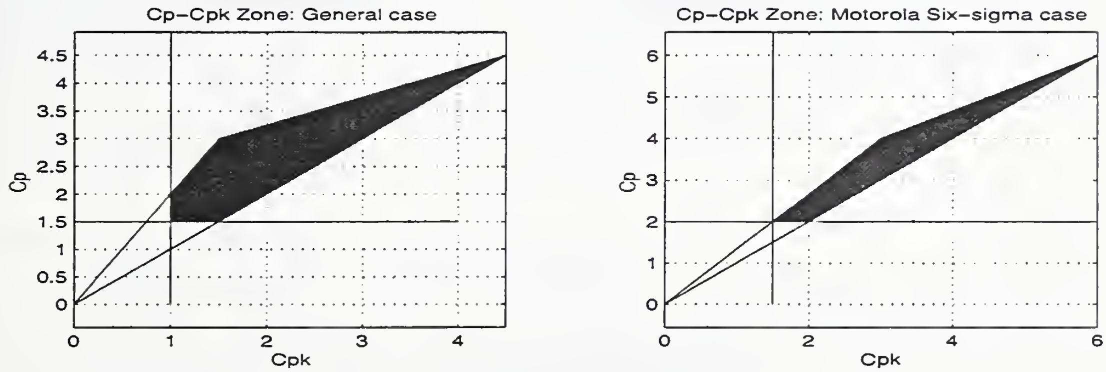


Figure 5: Tolerance zone in $C_{pk} - C_p$ plane

Zone in $\mu - \sigma$ Plane Using C_{pk}, C_p

The triplet (P, K, F) gives rise to a different shape to the above tolerance zone in the $(\mu - \sigma)$ plane. Since $\sigma \geq 0$, this tolerance zone is the intersection in the upper half-plane of the following half-planes.

$$\begin{aligned}
C_p &= \tfrac{USL-LSL}{6\sigma} \geq P & \Rightarrow \ & \sigma \leq \tfrac{USL-LSL}{6P} \\
C_{pk} &= \tfrac{\min\{USL-\mu, \mu-LSL\}}{3\sigma} \geq K & \Rightarrow \ & \sigma \leq \tfrac{\mu-LSL}{3K} \text{ and } \sigma \leq \tfrac{-\mu+USL}{3K} \\
C_c &= \tfrac{|\mu-MSL)|}{d} \leq F & \Rightarrow \ & MSL - Fd \leq \mu \leq MSL + Fd
\end{aligned}$$

where $MS = \frac{1}{2}(LSL + USL)$ and $d = \frac{1}{2}(USL - LSL)$. Figure 6 illustrates both the general and the six-sigma case. For the same triplet P, K, F, each point in the $\mu - \sigma$ tolerance zone maps to a point in the tolerance zone in the $C_{pk} - C_p$ plane; and each point in the $C_{pk} - C_p$ tolerance zone is the image of two points in the corresponding $\mu - \sigma$ tolerance zone (from the definition of C_{pk}).

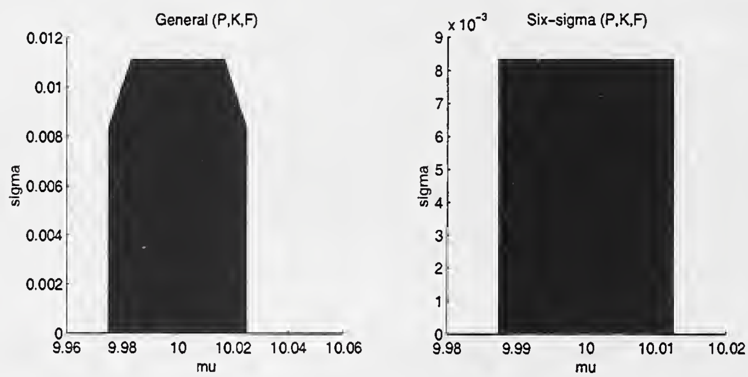


Figure 6: Tolerance zone in $\mu - \sigma$ plane

Zone in $\mu - \sigma$ Plane Using C_{pm}

We use the index C_{pm} to define the tolerance zone in $\mu - \sigma$ plane as follows:

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} \geq M \Rightarrow (\sigma^2 + (\mu - T)^2) \leq R^2$$

where $R^2 = \frac{(USL-LSL)^2}{36M^2}$. We get the tolerance zone as a semi-circle as shown in the Figure 7.

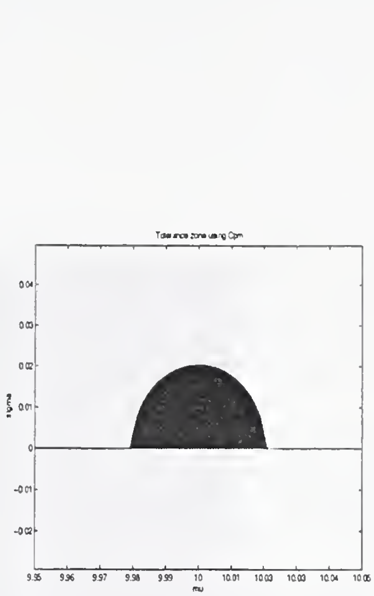


Figure 7: Tolerance zone in $\mu - \sigma$ plane using C_{pm}

Nonparametric Statistics

In the previous section, we discussed tolerance zones that are defined in the space of parameters of the population, namely the first two (central) moments of the distribution. In the nonparametric statistics, the distribution function zones are defined using the whole distribution. If the measured characteristic X is normally distributed and if we plot the cumulative distribution function (CDF) of X on normal probability paper, it will be a straight line. The slope of the line is $1/\sigma$ and the x-coordinate of the intersection of this line with $y = 0.5$ line is the mean (for normal distribution mean is equal to median). Each line with positive slope corresponds to a unique normal distribution.

Each point in the $\mu - \sigma$ plane corresponds to a line in the normal probability plot and a line segment in the $\mu - \sigma$ plane corresponds to a zone in the normal probability plot obtained by sweeping a line. If the line segment in $\mu - \sigma$ has a non-zero slope, then this sweep is a pure rotational sweep of a line in the normal probability paper. The horizontal line segment in $\mu - \sigma$ translates to a line in normal probability plot. For complete details refer to [72]. The collection of normal CDFs which belong to such a tolerance zone is independent of the choice of plotting in a normal probability paper. This collection is called the CDF tolerance zone [72]. Figure 8 shows a typical CDF tolerance zone.

Composition of Statistical Tolerance Zones

Until now we have discussed statistical tolerance zones for piece parts. Now, we extend this to an assembly level tolerancing. In order to compute the tolerance zones for Y we need to formulate the rules of composition of statistical tolerance zones, i.e., we need mathematical rules for combining part level variations to produce assembly level variations. Specifying individual part tolerances statistically as axis-parallel rectangular zones in the (μ, σ) plane will result in simple evaluation of the mathematical rule. In the general case, such as triangular

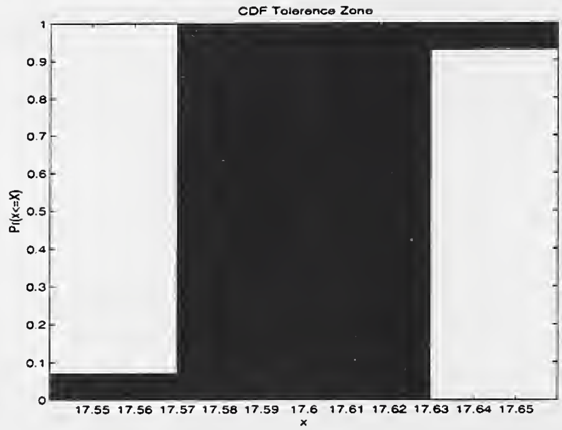


Figure 8: CDF Tolerance Zone

statistical tolerance zones, the rule of composition involves computing Minkowski sum [74] of objects bounded by curve segments even under the assumption of statistical independence. The rule of composition for CDF tolerance zones is totally different and involves advanced mathematical computations. The working group WG 13 of ISO TC213 is evaluating this approach.

Let us discuss the composition of statistical tolerance zones for a linear case of ARF $Y = a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n$, where $a_i's$ are some constants. We also assume that $X_i's$ are independent so that

$$\sigma_y^2 = \sum a_i^2\sigma_i^2 \tag{6}$$

where (μ_i, σ_i) are the mean and standard deviation of X_i respectively.

Our task is to compute the statistical tolerance zone for Y. To do this, in [74], they have defined statistical tolerance zone in $(\mu - \sigma^2)$ plane and use (6) to reduce the composition of statistical tolerance zone s to simple Minkowski sum. Refer [74] for more details.

Statistical Tolerance Zones: Summary

Given an arbitrary normal distribution for the variable X, we can check if the population is acceptable by checking any of the following: (i) if the tuple (C_{pk}, C_p) is within $C_{pk} - C_p$ zone, or (ii) if an ordered pair $(\mu - \sigma)$ is within the $\mu - \sigma$ zone, o (iii) if the CDF of X lies within the CDF zone. If $X \sim N(\mu, \sigma)$ then all the above discussed zones are equivalent; that is, the normal population is acceptable according to one interpretation, if and only if it is acceptable in the other two interpretations also [72]. In order to apply this methodology for general cases, like non-normal and unsymmetric, we need to generalize the above tolerance zone concepts. This may be necessary for the interpretation of statistical tolerancing when applied to geometric tolerances, such as perpendicularity. In this case we would ideally like

to produce cylindrical features with zero perpendicularity tolerance. This implies that the target should be LSL. We can enrich the class of distributions by dropping the normality assumption. This may be necessary, because perfect normality is seldom realized in practice. Even in theory the normal distribution might not be expected for certain tolerances (for example, perpendicularity).

If the process is centered and the target is the mid-point of the specification limits, then the indices are equivalent. If the last two assumptions are relaxed (that is unsymmetric and non-normal) care should be taken in interpreting the values of the indices. But as long as the distributions are determined by their first two moments these indices will serve as useful tool to assess the process performance. Figure 9 summarizes the various zones.

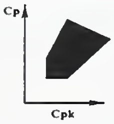
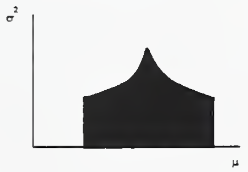
Parameter(s)	Tolerance Zone	Remarks
Parametric Statistics Cpk-Cp zone		More appealing to process and quality engineering. We need to specify (P,K,F)
Using Cp,Cpk (μ,σ) zone Using Cpm	 	Equivalent to the above. Easier to perform statistical tolerance analysis and synthesis. In case 1, we need to specify (P.K.F) and in case 2 we need to specify lower bound for Cpm
μ,σ^2 zone		Under the assumption that X's are independent the composition of ST Zones is given by simple Minkowski sum.
Nonparametric Statistics CDF zone		Greater flexibility and generality and can accomodate non-normal and unsymmetric situations. If the X ~ N(μ,σ) and process is statistical control, then all the zones are equivalent. For non-normal cases, as long as the distributions are determined by the first two (cenral) moments, first two zones agree exept for minor exeptions.

Figure 9: Various Tolerance Zones: Summary

3 Function-Assembly-Behavior Model

In this report our main aim is to develop an integrated comprehensive function, assembly (artifact) and behavior model. To do this, we need to develop appropriate assembly models and tolerance models that would enable the designer to incrementally understand the build-up or propagation of tolerances and optimize the layout, features, or assembly realizations. This will ensure ease of tolerance delivery. In [53], a multi-level approach called Design for Tolerance [DFT] process was proposed. The DFT enables tolerancing to be addressed at successive stages of design in an incremental fashion. In Section 4, we address the effective use of the FAB and DFT model for design tolerancing starting from the conceptual stage of the design and evolving continuously throughout the entire design process to the final detailed design.

3.1 Literature Review

Attempts to describe an assembly model in a CAD environment have been made by a number of researchers [21, 41, 62, 67, 70] in order to carry out assembly analysis, kinematics analysis, and tolerance analysis. Five major issues related to automated assembly have been studied elaborately. These are: 1) assembly, component and feature geometry, and topology representation; 2) identification and establishment of functional relationships between components of an assembly; 3) identification of precedence relationships; 4) assembly sequence/plan generation; and 5) assembly analysis. Since mating relations establish geometric constraints between components, several researchers have focussed on characterizing mating conditions and the impact on tolerance synthesis and analysis. Lee and Gossard [47] suggested the inclusion of mating information in a data structure representing an assembly for CAD/CAM systems. Work by Roy et al. [62, 67] suggests that an assembly can be decomposed into component features eventually; thus mating conditions among components can be represented as a relation among features by a 'modified functional relationship graph'. Srikanth et al. [70] investigated the mating relations among parts and developed a 'relation' operator to represent mating conditions in Relational Graph, which is a graph for representing an assembly. This representation can be applied for dimensional tolerance analysis of an assembly, but not for geometrical tolerances among features of the component parts. Rajan et al. [61] report the development of assembly representations that capture mating constraint information generated during the detailed design stage. Important mating constraints, joint attributes, and a relational structure are explicitly embedded within the hierarchical assembly structure to facilitate assembly analysis. The joint kinematics information and component degrees of freedom available from the assembly model can be used in formulating a mathematical model for tolerance analysis and synthesis.

Wang and Ozsoy [81] have described representation of assemblies for automatic generation of tolerance chains. The assembly is represented in an elaborate data structure with

information on assembly decomposition; (4x4) transformation matrix for each instance of a component/sub-assembly mating feature; mating conditions (against, parallel, fit), dimensions and tolerances of mating features, etc. A tolerance chain for any given assembly can be generated algorithmically using the above information. Tolerance analysis uses these chains. The representation does not need geometric data but cannot be used in early stages of design due to the nature of information required to complete the data structure.

In [80], a graph-based representation scheme-called the Tolerance Network (TN)- was proposed to represent the geometric dimensioning and tolerancing (GD&T) specifications in a feature-based product model. Tolerances are treated as variational geometric constraints and the authors explain how the tolerance network (symbolic constraint graph) represents the designer's intentions regarding functional, manufacturing, and quality control of the product. In their paper the authors have considered 'assembly graph' (an undirected graph with parts as nodes and connection constraints as arcs) as an assembly model for tolerancing. The TN can be a good candidate for tolerance representation.

The need for assembly model standards is one of the major suggestions from the participants of the national workshop at NIST [76]. The Japanese National Committee (JNC) has proposed a neutral assembly model for inclusion in STEP [77]. An extensive review of assembly requirements, representations, and methodologies that support assembly design and planning can be found in[61]. Pratt [51] has developed a more generic geometric constraint representation for use in describing component as well as assembly constraints.

With tolerance analysis as the main objective, Whitney, Gilbert, and Jastrzebski [82] proposed a model of assembly. This model contains the following information: mating features that build up the assembly; a graph representing the mating part; an underlying coordinate structure of the assembly; and homogeneous (4x4) matrix transformations to represent dimensions and tolerances of each part (in accordance with Y14.5M-1982 geometrical tolerancing standard). The transformations represent both nominal relations between parts and variations caused by geometric deviations allowed by tolerances. These transformations can be used to propagate tolerances through an assembly so as to compute the location of any designated part starting from a reference part, taking into account variations in the locations, sizes, and shapes. The above representation can potentially be used in the early design stage. The concept of Datum Flow Chain (DFC) advocated by Mantripragada and Whitney [49] can be used to provide information necessary for locating parts accurately with respect to each other, to relate the datum logic explicitly to product key characteristics, assembly sequences, and choice of mating features and to provide information needed for tolerance analysis.

Ge Qu [60] reported a method on automatic tolerance allocation (in a linear tolerance stack-up) with process selection based on hierarchical assembly model. In his work, an assembly is modeled by hierarchical data structure and assembly relationships between components are specified by mating relations through the mating surfaces of each component. Based on these mating relationships, the method can generate dimension graphs (to rep-

resent dimension vector loops in the assembly) where nodes correspond to mating surfaces and edges correspond to dimensions (and tolerances) between mating surfaces. The dimension graphs provide the basic mathematical models for solving tolerance allocation problems. This research work addresses linear tolerance stack-up issues for only three types of fit relations: length fits, cylindrical fits and housing-shafts. It uses a tree search-based technique for solving discrete tolerance and process selection problems in length fits and a rule-based method to specify tolerances for cylindrical fits.

What is missing from previous works is a general and unified representation scheme for geometric tolerances and mating conditions across parts and assemblies. In addition, very little consideration has been given to the significance of an assembly data model for pro-active analysis of the product (especially for tolerance synthesis and analysis) as design progresses. In this paper, we discuss the integration of function, assembly, and behavior representation into a comprehensive information model (FAB models). To do this, we need to develop appropriate assembly models and tolerance models that enable the designer to incrementally understand the build-up or propagation of tolerances and optimize the layout, features, or assembly realizations. This will ensure ease of tolerance delivery.

3.2 Information Requirements for Synthesis and Analysis of Tolerances

Significant productivity gains can be achieved if we consider tolerance information through out the entire product design cycle. We need to develop appropriate assembly models and tolerance models that would enable the designer to incrementally understand the build-up or propagation of tolerances. We also need to optimize the layout, features, or assembly realizations so as to ensure ease of tolerance delivery.

Any valid assembly sequence has a direct influence on the synthesis of tolerances. At the same time, it should be kept in mind that any tolerances specified also directly influence the manufacturing and assembly tasks. The tolerance synthesis, analysis, and assembly planning tasks cannot be treated separately; they are dependent on each other. Therefore, we need to derive some kind of compromise between the tolerance design and the various planning tasks. This compromise should not violate the established rules and procedures in assembly planning. It should be able to take account of the lack of tolerance information in the conceptual stage of assembly design and in the formulation of various plans. Different types of information are needed at different abstract levels of assembly and tolerancing. At the conceptual level, with a view towards computerized tolerance synthesis and analysis, we need the following types of information: a description and quantification of the functional requirements of the product-the design intent; information that leads to the selection of assembly configuration (layout); identification of datums, features and geometric relations between features that will influence functional and assembly requirements; and an assembly

graph that provides information for the selection of location logic (in deciding the manner in which parts are located with respect to one another).

During tolerance synthesis and analysis, an assembly model should provide information regarding representation of assembly design in terms of attributes such as design function, design parameters, geometric characteristics, manufacturing features and processes, and assembly conditions [63]. The decomposition of overall part functional requirements (PFRs) and their mapping into different product specifications are important in the context of product/assembly design particularly during synthesis of geometry and tolerances. The most important information that is essential for tolerance synthesis is the geometric characteristics of each component and inter-relationships between different components in an assembly. The geometric characteristics of a component consist of size (geometric dimensions), form (type of geometry, i.e. planar or cylindrical surfaces) and profile (nature of surface) while inter-relationships consist of relative positioning of components.

For tolerance analysis in the detailed engineering stage, an assembly model should provide kinematics inputs for assembled components that are small variations of component dimensions around their nominal values, and detailed constraint information on the interaction between mating parts. These constraints also serve as functions-assembly response function (ARF)-by which assembly variation may be studied. At this stage, the assembly model provides information on dimension vectors that is used to represent various dimensions. Each vector represents either a component dimension or kinematically variable dimension that can be arranged in chains or loops and represents those dimensions that stack together to determine resulting assembly dimensions. The other model elements include datum reference frames, assembly and component tolerance specifications, feature datums and geometric feature tolerances, and kinematics joints that describe motion constraints at the points of contact between mating parts. It should also provide models of assembly variations that include 1) dimensional variations, 2) geometric variations, and 3) kinematics variations.

The assembly model introduces more than just assembly (and manufacturing) information for tolerance analysis and synthesis. It serves other functions, such as the functional analysis of the whole assembly, and assembly sequence issues. The assembly model provides detailed knowledge/information of part geometry and function.

3.3 Information Models for the Design for Tolerance

In this section, we will first describe the representational issues of functional models and the representation of the FAB model. Then, we will discuss part models in the assembly and applicable tolerance models for both the part and assembly.

In the early design phase, most of the design decisions taken are concerned with the desired characteristics and the overall functions of the assembly. By function we mean an abstract formulation (or definition) of a task that is independent of any particular solution. In this phase, the abstract functional specification of an assembly is transformed into a

physical description. In the later design phases, the physical decisions that are made in the earlier phases are elaborated to ensure that they satisfy the specified functional requirements and life cycle evaluation criteria. In order to manipulate the function information, a *functional data model*, that describes the functional information throughout the design cycle, is needed. Appropriate reasoning modules can then interoperate and extract functional information during the decision-making processes from the functional data model (as geometric reasoning modules query data from the *product data model*, that is the CAD model, for shape information).

3.3.1 Function Representation and Data Model

Depending on the design phase, different types of functions exist at different levels of abstraction. In the conceptual design phase, functions are usually independent of the product domain, whereas in later design phases, when the functions are detailed, they become increasingly dependent on the product domain that has been selected. Therefore, we need to adopt an appropriate formal function representation scheme that will be helpful when modeling the overall function of the assembly in the conceptual design stage, and will also be useful to model a smaller sub-assembly, particularly as component and feature levels are approached. The representation should also have unambiguous semantics in order to perform analysis on (i.e. to manipulate) the functional description. Many researchers [5, 12, 44] have already investigated different types of functional representation schemes. Among these, two types of representation are particularly relevant in the domain of electro-mechanical products: 1) model of the flows between inputs and outputs of a product, and 2) syntactic languages, which use a grammar consisting of verbs and nouns to describe a product. Refer to [5, 12, 35, 44, 49] for a detailed discussion on these techniques. Both representations have limited use in the synthesis of mechanical devices. They are not able to support applications over the complete product life cycle. Our functional data model can represent the functional description of a product through out the entire product design cycle.

For computer representation of function and behavior, we adopt the object-oriented design modeling language that has been developed as part of the Design Repository Project at the National Institute of Standards and Technology [52]. A **Function** has a set of inputs and outputs, and may have other attributes or relationships that specify sub-functions and composite functions. Functions are treated as the instantiations of classes of functions. The **Function** hierarchy consists of four subclasses of function: **Transform_function, Convey_function, Supply_function and Control_funcion** (for a detailed description, refer [52]). Inputs and Outputs of functions are described in terms of sets of instances of a subclass of the general **Fluent** class. Fluents include both physical and abstract phenomena that are associated with inputs and outputs to functions, such as motion, force, heat, liquids, current, and so on. The **Fluent** class is divided into two main subclasses: **Abstract_fluent** and **Physical_fluent**. The **Abstract_fluent** class specifies things like motion and electric-

ity while **Physical_fluent** class specify things like gases and liquids. All **Fluent** classes have slots for explicit references, to which artifact the fluent is flowing from, and where it is flowing to. The set of subclasses for the general **Function** and **Fluent** classes together describe the semantics of the functional and behavioral aspect of the object-oriented design modeling language.

Function-Assembly-Behavior Model

Figure 10 depicts a typical FAB data model for capturing the product development-related issues from the conceptual design stage to the detailed assembly building process. The main purpose of the proposed aggregate structure of the function, behavior and assembly together in this data model is to support conceptual design as well as design for manufacturing and assembly, starting from an early design stage. It starts with defining the design context that represents a particular design domain and consists of the design tasks (i.e., the Goal) relevant to the current design assignment, the decisions that have been made, the design requirement description and the artifact that has been developed as part of the design process. The design requirements have three attributes: `description, weight and dictated_by`. The `description` records the requirement description that covers input specifications by the customer, objectives, and specified constraints. Following a formal method like Quality Function Deployment (QFD), customer requirements, stake-holder requirements, operational requirements, and other statutory requirements are developed. These requirements can be further categorized into: 1) spoken expected requirements (i.e. directly specified by the customer); 2) unspoken expected requirements (i.e. automatically expected by the customer); and 3) unspoken unexpected requirements (i.e. no direct need but it will be nice to have them). The `weight` attribute describes whether the requirement is a *demand* or *wish.* The product requirements are finally `fulfilled_by` a set of functions.

The **Function** (a function means what the system/product is for) is associated with the transformation of an input physical entity into an output physical entity by a part. A function has the attributes: `name`, `inputs`, `outputs`, `dictated_by`, and `achieved_by`. The transformation between the physical entities (or fluents in inputs and outputs) is controlled by the behavior of the **Artifact** (i.e., assembly). This, in turn is governed by part/sub-assembly/assembly geometry, either by itself or in association with other parts/sub-assemblies in an assembly, and the physical laws which are associated with those physical entities. The `fulfilled_by` attribute of the function relates it with a (or a set of) product requirements. The `fulfilled_by` and `dictated_by` relationships between product requirements and functions thus facilitate both top-down and bottom-up of design synthesis processes. An **Artifact** which can either be an assembly, sub-assembly, individual part, feature, or even any low-level geometric entity (i.e., face, edge, or vertex) achieves the **Function**.

The assembly/part behavior is achieved through certain functional relationships associated with its sub-assembly/features [20]. In the conceptual design phase, each functional

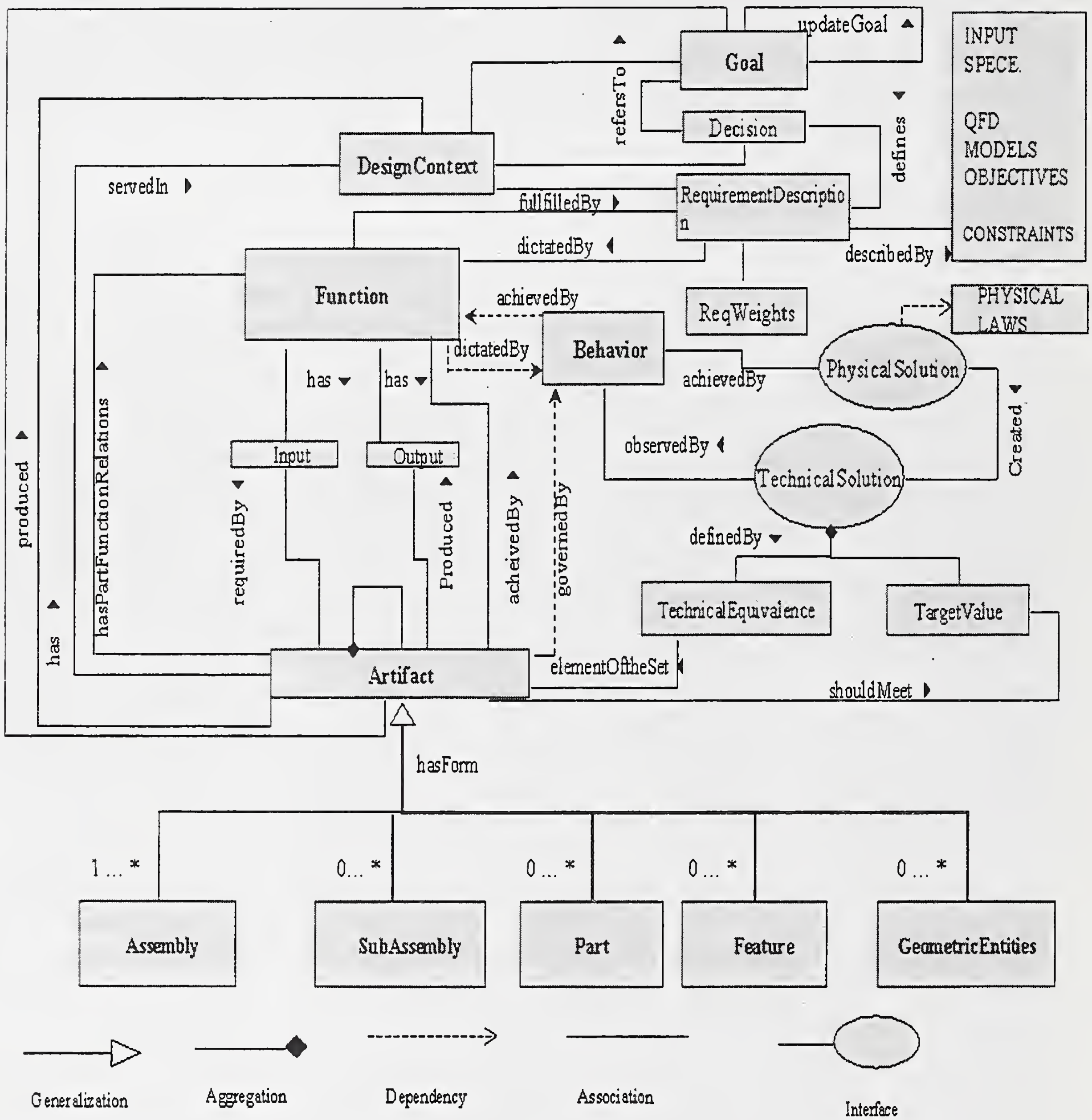


Figure 10: Function Assembly Behavior Model

relationship is translated to its spatial equivalents using appropriate physical laws. A part

achieves its behavior through the interactions of its geometry or environment. Thus, the utility of part behavior can only be realized when it is decomposed into an equivalent part-function model. The design solutions are obtained as a composition of physical solutions and technical solutions. A physical solution is the first step to realize behavioral requirements and is governed by physical laws. The technical solutions are technical manifestations of the physical solution. A technical equivalence (i.e. a physical product) determines the final product/artifact and defines final *target values* that are expected to achieve the desired function.

The assembly data model contains assembly-specific information of a product that not only helps in detailed assembly analysis (i.e., determining assembly sequence, etc.) but also in synthesizing the system's design and studying its behavior. Its attributes are: `key assembly characteristics`, `required assembly equipment and accessories`, `sub-assembly/part`, `assembly tolerances` and `assembly kinematics variations`. The product's important geometric features and material property information (directly derived from the customer requirements) that are highly constrained are given by `key assembly characteristics` [49]. These assembly key characteristics cannot be violated in any situation to achieve a desired level of the product's performance, function, and form. The description and specification of required assembly equipment and accessories are necessary for the assembly analysis at the detailed design level. The sub-assembly/part defines the composition of an assembly. Assembly tolerances are applied to the assembly dimensions which are kinematics assembly variables, that arise during the assembly of components. It should be noted that while component dimensions are subjected to random process variations, the assembly dimensions are not manufacturing process variables; these are kinematics assembly variables. Kinematics variations occur at assembly time, whenever small adjustments between mating parts are required to accommodate dimensional or form variations.

The sub-assembly/part has five major attributes: `name`, `identification number`, `a set of features`, `relationship with other sub-assembly or part` and `a pointer` to its parts' CAD database (that supplies geometrical, topological, variational (tolerance) and other technical information). The parts' CAD database contains detailed constructive solid geometry/boundary representation (CSG/B-Rep) information. A feature is composed of feature identification number, feature datum, feature locations and orientation, geometric feature tolerance, process characteristics for manufacturing the feature, and its spatial relation with other feature(s).

Since dimensional variations (lengths, angles, etc.), geometric form, and feature variations (flatness, roundness, angularity, etc.) are the result of variations in the manufacturing processes, the process characteristics hold key attributes. The spatial relation of a feature with other feature(s) describes 1) datum reference frames (DRFs), 2) type of relationships among features (i.e., feature interaction list), and 3) and any other feature characteristics (e.g., face pointers to the respective B-Reps of the CAD database). The spatial relations (with other feature(s)) may include constraints such as articulation representation for electro-mechanical

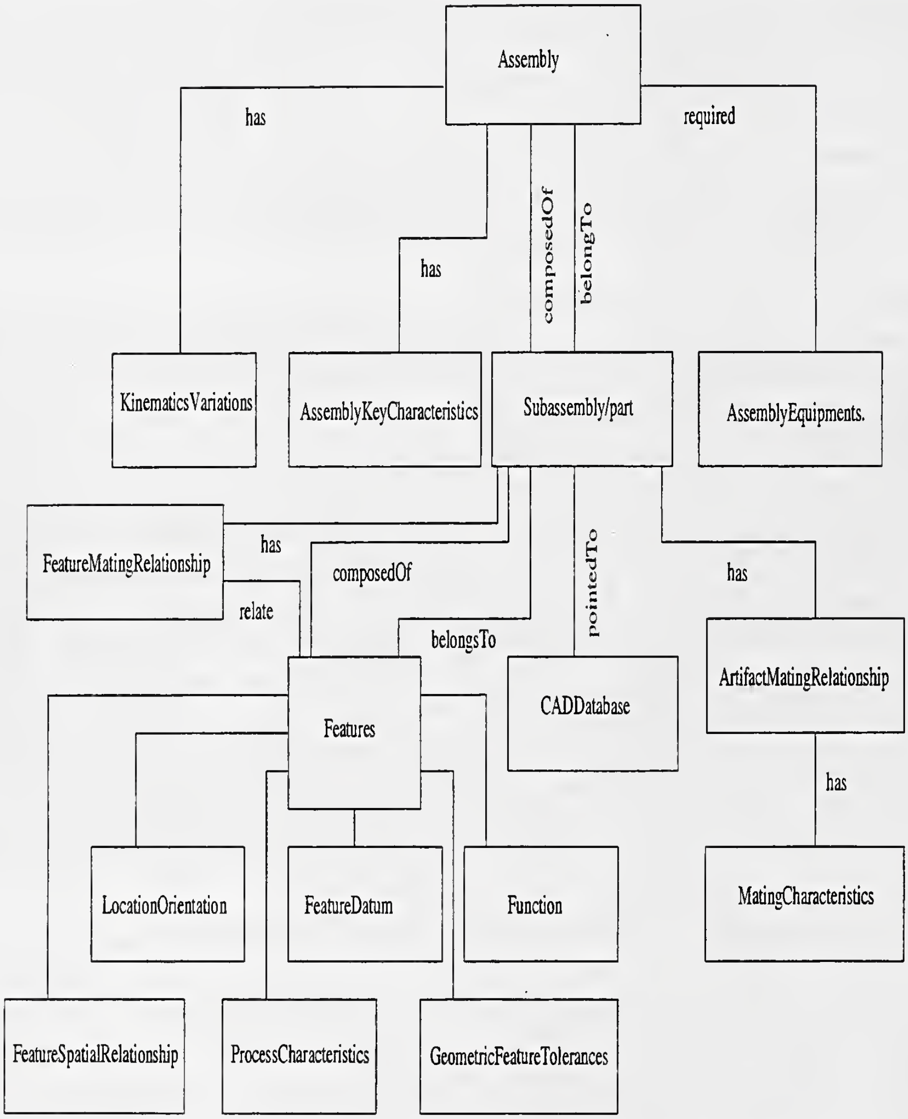


Figure 11: Function Assembly Behavior Model:Assembly

part at the feature level (e.g., position coincidence, rotation axis coincidence, parallel normals, center coincidence)

In the sub-assembly/part level the relationship of a sub-assembly/part with another sub-assembly/part has two main attributes: relationship type and mating characteristics. There are various relation types: `contact`, `detachable type rigid attachment`, `constraint` etc. Mating characteristics are described by: 1) mating sub-assembly/part identification number,

2) identification numbers of mating features, and 3) mating conditions. Each face contains the following information about face-type (planar, cylindrical, etc.):

- `dimension`,
- `size_tolerance`,
- `form_tolerance`,
- `normal vector`,
- `process model`, `force_d.o.f`,
- `kinematic_d.o.f`,
- `surface_roughness`, and
- `face_interaction_list`, refer [7] for details.

Mating conditions are described by mating type, joint characteristics, spatial constraints, rules, and attributes. Mating type can be any of the following: fit (i.e., sliding, clearance, transition, and interference),fasteners, gear, bearing, etc. These can be further described in terms of some basic relationships: `orientation`, `primary mating relations` (no_contact, touch_contact and overlap_contact), `size relation`, `intersects`, `between`, `contains`, and `connect`.

In this representation, assembly joints are defined by two joint attributes: rigid and movable [61]. Rigid joints require only the position (location and orientation) information to be specified, whereas movable joints require the specification of position and kinematics information. Kinematics information specifies the type of joint (revolute, cylindrical, prismatic, spherical, etc.). Note that this information is explicitly embedded in the assembly data model for the purpose of information completeness. This information is directly usable in the detailed assembly analysis phase. Knowing the contact relationships between various parts would facilitate the determination of joint characteristics in an automated manner.

Similarly, spatial constraints may also be explicitly specified or may be implicitly specified by the adjacency of component surfaces. Depending on the nature and degrees-of-freedom, spatial constraints can be of different types: completely unconstrained, single planar, multiple planar, planar motion, constrained planar motion, cylindrical, uni-cylindrical, prismatic, uni-prismatic, revolute, and completely constrained (refer to [61] for a detailed discussion).

Rules and attributes represent other necessary technical information about the mating condition. The technical information describes properties of the joint connection and restrictions that are applied in the connection between parts/sub-assemblies. For example, in a rivet joint, the strength of the rivet joint is described by attributes and rules of the joint connection.

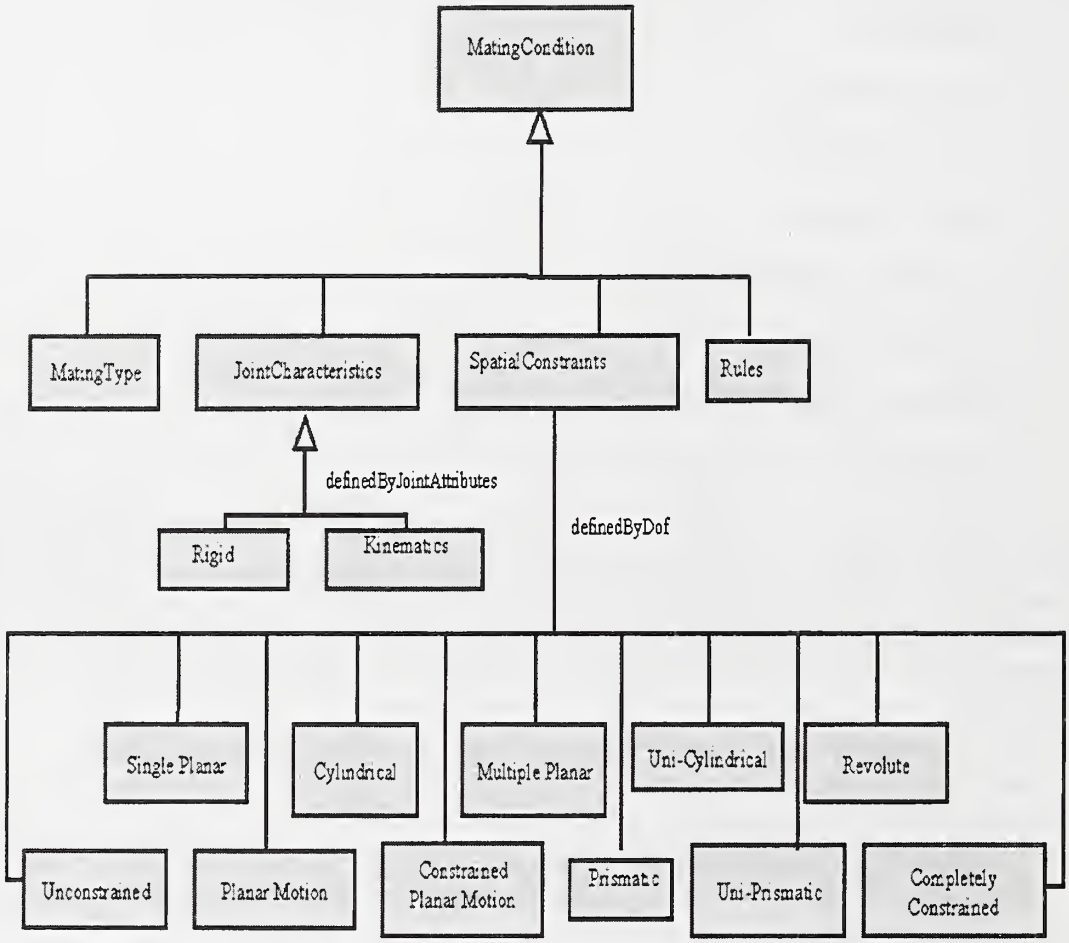


Figure 12: Function Assembly Behavior Model - Mating Characteristics

The Function-Assembly-Behavior data model described so far provides information for the analysis of product life-cycle process. Not all the information is available in the early conceptual design. The information content of the data model evolves with the product design process.

The complete object oriented class definitions for Function, Artifact and Behavior are under development and will be reported in a future NIST publication.

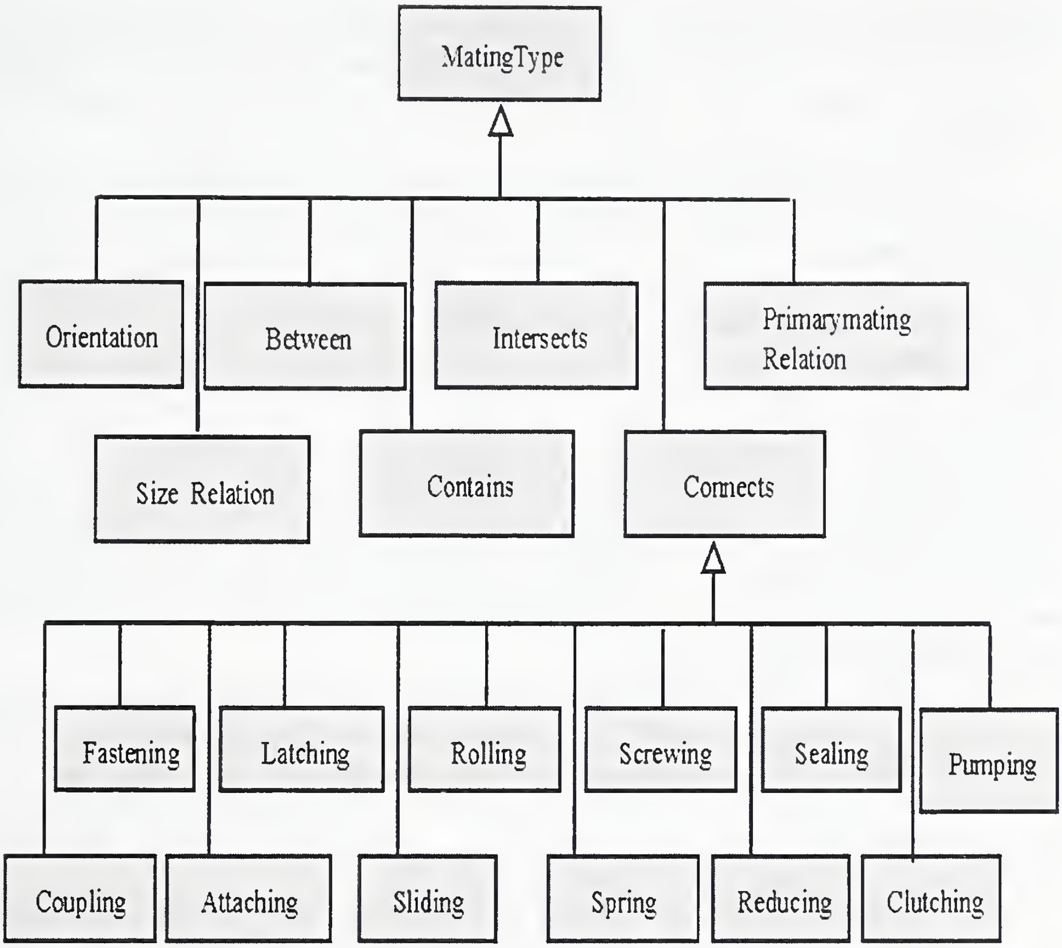


Figure 13: Function Assembly Behavior Model- Mating Types

4 FAB-DFT Integration

In the previous section we discussed and explained the object oriented architecture of the FAB model. In this section we discuss how a designer can use this FAB model for design tolerancing. In [53] a dynamic architecture, called Design For Tolerancing (DFT), has been proposed to help the designers to make tolerance-related decisions from the conceptual design to detailed design. The basic motivation of the DFT architecture is the definition of a multi-

level approach that enables tolerancing to be addressed at successive stages of design in an incremental fashion. The resulting design process integrates three important domains: (1) design activities at successive stages of design; (2) assembly models that evolve continuously through the design process; and (3) methods and best practices for tolerance analysis and synthesis.

The basic motivation for defining the FAB model is to change the traditional thinking of design as form-to-function transformation. Present CAD systems have a wealth of tools for generating geometric forms for the objects. This forces the designers to decide on the form first then think about the function later. This way of designing has the potential problem of designing sub-optimal designs that are a mere modifications of existing ones. Present CAD systems cannot help the designers to be innovative. Hence, tools for helping a designer to think in terms of function should be developed and form should result from function, i.e., function-to-form.

Knowledge-based design systems implement this paradigm by first focusing on the symbolic aspects of design and later mapping the symbolic structure to a geometric model. They can also capture the various semantic relationships between design objects. Essentially, knowledge-based systems use techniques developed by artificial intelligence researchers to capture the knowledge of expert designers in a computer. These systems help the engineer in the design generation process.

The proposed FAB model goes one step further to include behavior modeling aspects into function-to-form method of design. The various classes defined in the object oriented model of the FAB will help a designer to study design activity, providing various viewpoints. For example, a designer can study assemblability analysis, ease of tolerance delivery, quality and supply chain etc. Our main focus is on tolerancing. The three threads of the DFT architecture are explained in Figure 14, source [53].

The FAB model evolves the function and form (assembly) for the entire design life cycle. The FAB-DFT integration is better explained through an example. Before we discuss the example, we discuss briefly about various research efforts on assembly process models.

In the literature, several researchers have presented their viewpoint of what the assembly design process should be. In [53] a brief outline of some viewpoints that emphasize tolerancing is provided using the SIMA (Systems Integration for Manufacturing Applications) reference architecture formulated at the National Institute of Standards and Technology [3], which provides a generic specification of design-related activities for electro-mechanical products. Figure 4 shows the various design stages and activities in the SIMA reference architecture. Stage A11 (*Plan Products*) involves developing the idea for the assembly depending on market needs and customer requirements and characterizing it in terms of function, target price range, and relationship to existing product lines. In Stage A12 (*Generate Product Specifications*), an engineering specification for the assembly is formulated. This involves mapping the customer requirements into engineering requirements and refining these in consideration of the relevant laws, regulations, patents, and product standards, etc. In Stage

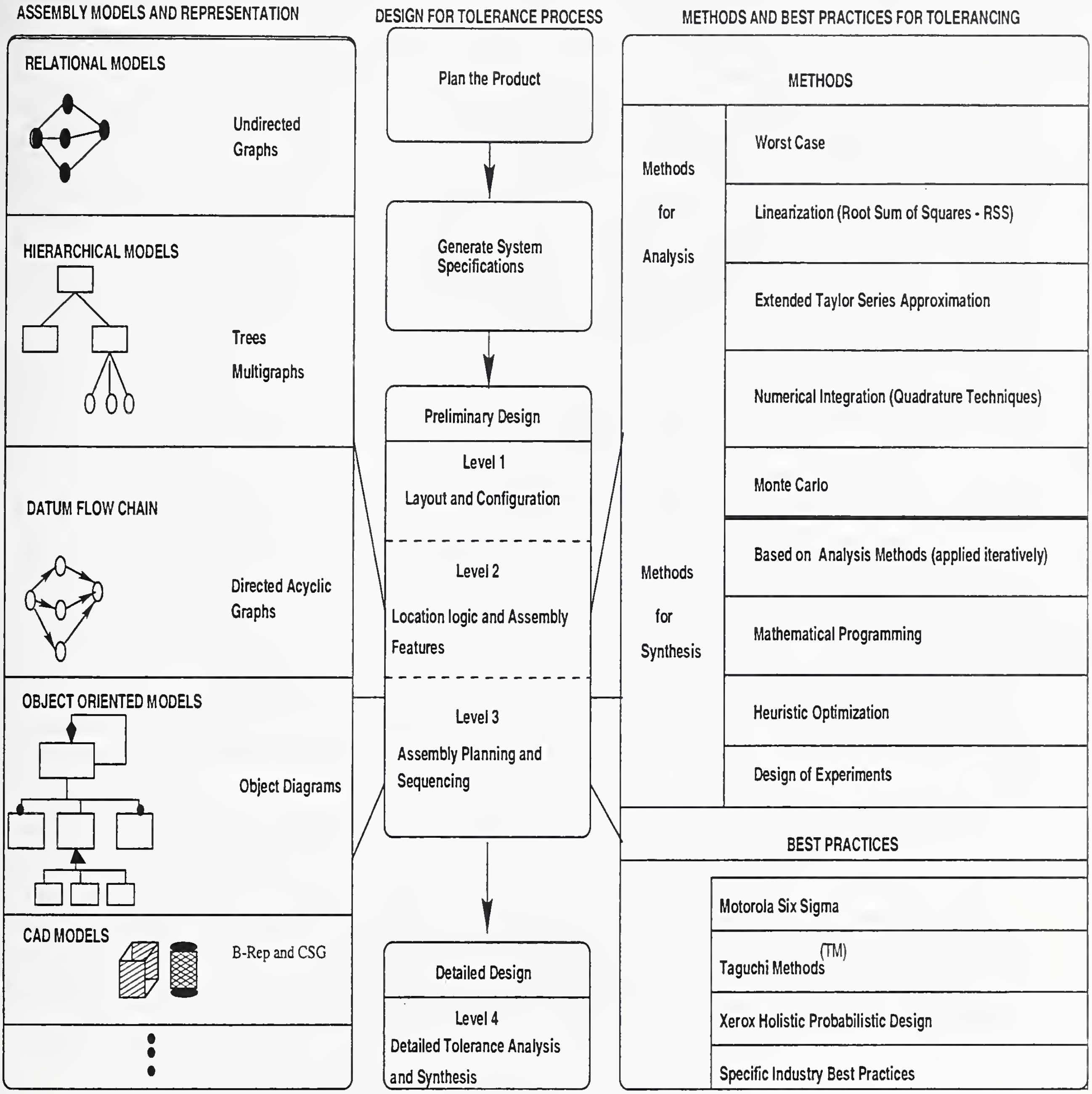


Figure 14: Assembly models, design process stages, and tolerancing tools

A1 : **DESIGN PRODUCT**

A11 : **Plan Products**

A12 : **Generate Product Specifications**

A13 : **Perform Preliminary Design**

A131 : Develop Functional Decompositions
A132 : Evaluate and Select Decompositions
A133 : Develop Preliminary Configurations
A134 : Consolidate Configurations
A135 : Evaluate Alternative Designs
A136 : Select Design

A14 : **Produce Detailed Designs**

A141 : Design System/Component
A142 : Analyze System/Component
A143 : Evaluate System/Component Design
A144 : Optimize Designs
A145 : Produce Assembly Drawings
A146 : Finalize System/Component Design

Figure 15: Design stages and activities in the SIMA reference architecture. Source: [3]

A13 (*Perform Preliminary Design*), the assembly design problem is decomposed into a set of component/subassembly design problems and specifications are developed for each component/subassembly problem. Interface specifications between the components/sub-assemblies are developed and a preliminary assembly layout is created. Finally, in Stage A14 (*Produce Detailed Designs*), all specifications needed to completely describe each subassembly or component are produced. This includes drawings and geometry, materials, finish requirements, assembly drawings, and fit and tolerance requirements.

4.1 Design Tolerancing: An Incremental Process

Potentially, tolerance considerations can influence the decisions taken at different design stages, in increasing level of detail. Also, the decisions taken at a particular stage influence

and can simplify the decisions taken in the downstream stages. Like other attributes of a product design, tolerance information changes over time, through successive stages from product planning to detailed design through on-going production. Hence a robust tolerance representation would be mutable and directly related to the evolving product design. The incremental refinement of design tolerancing processes and tolerance representations proceeds in symbiotic fashion. Consider, for example, a tooling design/build process. Both lead time and cost for tooling is often highly dependent on the tightness of the tolerance requirements. Scheduling of rough cutting for a die or mold can typically proceed prior to a final tolerance specification, but the finish cut, polishing, etc. must proceed afterward. Conversely, tolerance specification for a snap-fit in a high-precision injection-molded part must be preceded by a decision about assembly process (e.g., manual or robotic). For complex assemblies with many parts, the timing and precedence of tolerancing decisions can profoundly affect scheduling and total lead time. Analysis and synthesis for critical tolerance stack-ups is clearly related to process plan refinements. There are opportunities to compress cycle time by improved modeling prior to detailed design, but compatible, incrementally-refined representations of tolerances and manufacturing processes are the key.

The incremental and continuous, ongoing nature of tolerance decision making enables a natural aggregation/decomposition of tolerancing activities as the design matures. Another way of viewing this is in terms of the pruning that this causes at successive stages in the space of feasible solutions to the design problem. Early on in the design process, the solution space has a staggering cardinality and tolerancing decisions, if taken in a continuous ongoing fashion, can lead to substantial early reduction in the space of possible solutions thus making the design process efficient. Another alternative view is in terms of marked reduction in design iterations or design rework. In this sense, design for tolerance is similar in spirit to design for manufacturing/assembly [9] that also has the effect of dramatically shrinking the space of solutions and reducing iterations. Furthermore, DFA, DFM, or such other design related strategies may have close coupling with tolerance-related decisions and may both influence and be influenced by tolerancing at various stages.

4.2 Design for Tolerance: A Multilevel Approach

The first two stages A11 and A12 of the SIMA reference architecture and also the early stages of other assembly design processes essentially involve mapping customer requirements into product ideas and product specifications. Tolerancing is not directly involved in these early stages, except in very abstract terms; however, these stages provide critical inputs to the tolerancing decisions in the rest of the design process.

Thus we focus on Stage A13 (*Perform Preliminary Design*) and Stage A14 (*Produce Detailed Designs*) of the SIMA reference architecture. We divide these stages into the following four *tolerance-related levels (TR Level)* and develop a four-level approach to design tolerancing. Note the difference between SIMA stages and tolerance-related levels here.

- SIMA Stage A13: **Perform Preliminary Design**
  - TR Level 1: Assembly Layout and Configuration
  - TR Level 2: Location Logic and Assembly Features
  - TR Level 3: Assembly Planning and Sequencing
- SIMA Stage A14: **Produce Detailed Designs**
  - TR Level 4: Detailed Tolerance Analysis and Synthesis

These levels are fairly representative and generic for electro-mechanical assemblies. Neither the number of levels nor the description of the individual levels is to be viewed as being definitive. Figure 16 captures the essence of this architecture for DFT.

4.2.1 TR Level 1: Assembly Layout and Configuration

Once the product concept is known and engineering specifications are generated based on the key characteristics, TR Level 1 of the proposed process can commence. TR Level 1 involves decisions regarding the preliminary assembly layout/configuration. Such decisions may include: rough allocation of space, number of sub-assemblies, the configuration of critical sub-assemblies, grouping of components into sub-assemblies, and rough layout of the assembly.

Using the FAB model we need to generate the following information to make tolerance-related decisions. They are:

1. rough shapes/form for the parts/sub-assemblies
2. parts list
3. parts location
4. layouts and configurations

The information thus generated can be described in the form of a liaison diagram (relations between parts or sub-assemblies), a tree (assembly decomposition), and a partial DFC (to capture whatever location logic is known at this point). Candidate layouts or configurations can be identified and represented using these models. These layouts or configurations and related manufacturing process selection typically might differ in terms of ease of tolerancing. The tolerancing considerations here are at a coarse level and may be directly influenced by customer specifications. To effect such high level tolerancing decisions, aggregate level manufacturing process capability data will be required and is often available at

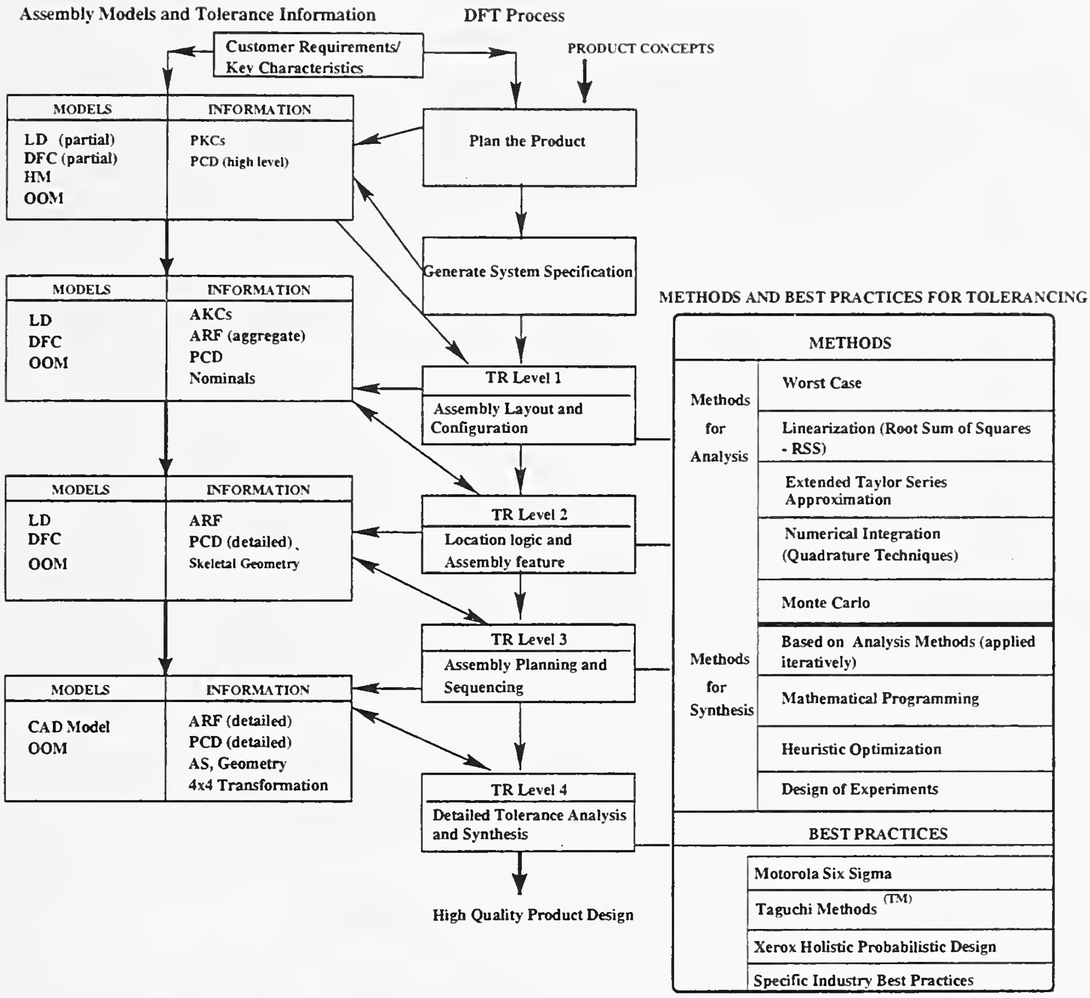


Figure 16: An architecture for design for tolerance. Legend: LD - Liaison Diagram; DFC - Datum Flow Chain; HM - Hierarchical Model; OOM - Object Oriented Model; AS - Assembly Sequence; PKCs - Product Key Characteristics; AKCs - Assembly Key Characteristics; ARF - Assembly Response Function; PCD - Process Capability Data

this point. Simple statistical assumptions and probabilistic calculations can be used at this stage. Also, for problems such as manufacturing process selection, the formulations such as in Section 2.2.3 can be used.

4.2.2 TR Level 2: Location Logic and Assembly Features

At the next level (TR Level 2), the following information is generated using the FAB model: assembly response functions (approximate); tolerance requirements at interfaces between major sub-assemblies and parts. The relevant process capability data could also made available. The decisions here are concerned with the location logic (how to locate sub-assemblies and components with respect to one another) and with choosing the appropriate assembly features to go with the location logic. The choice of features itself might depend on the assembly sequence (not the detailed sequence but a precedence specification among major assembly steps). The DFC model is suitable to capture the available/evolving assembly information here. There is close coupling among selection of features, selection of assembly sequence, and creation of DFC. Assembly models such as liaison diagrams are also relevant here. If the assembly is of Type 1, then the assembly features are predominantly decided by the functional requirements; if the assembly is of type 2, then the choice of assembly features is an important problem by itself. In the latter case, the DFC alone will not be adequate to conduct a tolerance analysis. A more detailed model that captures the tolerance flow at this level will be required. Tolerance analysis here can tell us which location logic is better from a tolerancing viewpoint and which set of assembly features would best accomplish tolerance achievement. This stage might also help us to find preliminary target values and tolerances for individual parts.

Statistical tolerance analysis methods mentioned in the preceding sections are all relevant here. Determining robust nominal values and preliminary settings of tolerances can be accomplished using Taguchi methods or Xerox HPD methodology.

4.2.3 TR Level 3: Assembly Planning and Sequencing

We proceed next to TR Level 3 where the detailed assembly response function, skeletal geometry of the assembly, assembly features, and specification of parametric or geometric tolerances of individual parts and features are again generated using the FAB model. The PCD could be made available (if applicable). From the tolerance specification, one may derive (4 x 4) matrix transforms for the nominals and variabilities associated with the parts [82]. The decisions here could be with respect to the selection of the detailed assembly sequence that achieves the required tolerance specifications in the best possible way. The models that we employed in the previous stage, like DFC and liaison diagrams, can again be used here. In fact, they are now updated with richer and more detailed information. This kind of representation and analysis is presented in [81], where several data structures

to capture tolerance-related information are presented. With the information available here, one can also carry out tolerance synthesis.

4.2.4 TR Level 4: Detailed Tolerance Analysis and Synthesis

TR Level 4 corresponds to the detailed assembly design stage. Here, the complete assembly sequence is known; geometric data about the parts and features is available; detailed part level tolerance requirements are known; the assembly response function is available in complete form; and low level process capability data is accessible. Detailed tolerance analysis and synthesis can be carried out here. Most tolerancing studies and tolerancing tools available support this level of design.

5 Current and Evolving Standards: Useful Inputs from FAB-DFT

The FAB model presented in this report can eventually lead to standards for tolerance and assembly representation and design process and product integration. The DFT and FAB integrated model will help the designers to make tolerance-related decisions from the conceptual design to detailed design and throughout the entire design life cycle.

The FAB and DFT framework uses both the top-down and bottom-up approach to the product design. The hierarchical assembly model defined in this report uses features as the atomic objects. The FAB model is defined in such a way that the various definitions and concepts that are already in use (ISO 10303-47 [26], ASME Y 14.5 [1,2]) and current evolving standards can be easily incorporated into the model. The FAB model can help the standards community in specifying the representation for assembly models, tolerance representation, and assembly level tolerancing. The class structure (**Function, Behavior, Artifact**) defined in FAB will help in unifying the various current and evolving definitions, concepts, and standards. In the following sections we briefly discuss ISO 10303-47 and the Geometric Product Specifications (GPS) standards being developed by ISO TC 213.

5.1 ISO 10303-47: Shape variation tolerances

This part of ISO 10303, [26], supports the dimensioning and tolerancing methods defined by the ISO 1101 [31] family of standards for the dimensions and tolerances on the engineering drawings. These methods include explicit dimensioning, associative dimensioning, plus-minus tolerancing, and geometric tolerancing. This part captures the semantic contents of the dimensions and tolerances as specified in those standards and applies to 3-D geometric models of parts.

The major subdivisions of this part of 10303 are:

- shape aspect definition: provides resources for the representation of shapes to which dimensions and tolerances are applied;
- shape dimension: provides resources for the representation of size and relative location to meet the dimensioning requirements found in the engineering design;
- shape tolerance: provides resources for the representation of limits within which manufactured shapes are permitted to vary.

However, ISO 10303 does not cover the following:

- the definition of the fundamental principles, concepts and terminology of tolerancing and dimensioning;
- the mathematical definition of tolerances and datums;
- the description of dimensioning or tolerancing practices;
- the specification of dimensioning inspection methods;
- the synthesis and analysis of tolerances;
- the tolerancing of product characteristics other than shape;
- the presentation of tolerances on engineering drawings;
- the specification of surface finish or surface roughness.

5.2 ISO TC213: Geometrical Product Specifications (GPS)

ISO TC 213's scope is standardization in the field of geometrical product specifications (GPS) i.e. macro- and micro-geometry specifications covering dimensional and geometrical tolerancing, surface with properties and the related verification principles, measuring equipment and calibration requirements including the uncertainty of dimensional and geometrical measurements. The standardization includes the basic layout and explanation of drawing indications (symbols).

ISO TC 213 is adopting a bottom-up approach for design. That is starting from feature level to the product level. The STEP community is adopting a top-down approach, that is, product-assembly-parts-features-conditions. These two approaches should logically meet at some common level.

ISO 5459 [28–30] is prepared by ISO TC 213. ISO 5459 cancels and replaces ISO 5459:1981, and a technical revision of ISO 5449:1981. It consists of the following parts, under the general title Geometrical product specifications (GPS) - Datums for geometrical tolerancing:

Part 1: Terms and definitions

Part 2: Datum and datum systems: Drawing indication

Part 3: Association methods for datums and datum systems for the assessment of geometrical tolerances

Part 4: Metrological establishment of datums and datum systems for the assessment of geometrical tolerances

The main definitions related to 5459-1 and 5459-2 are described in ISO/CD 17450: Geometrical Product specifications (GPS)-Model for geometric specification and verification [32]. The basic mathematical idea is as follows: Given a set of features, there are some operations (partition, extraction, collection etc.) by which we can get some other features that belong to this set. This way we can build the part, sub-assembly, assembly, and finally the product. This is the bottom-up approach taht was mentioned earlier in the report. The second part of ISO 5459 [29] describes the rules, explanations and manner in which datums, and datum systems are indicated in technical product specification. This part defines Single datums, common datums and datum systems. It also defines the writing and reading rules of the graphical language (drawing indications). Currently this rule based system needs a mathematical rigor.

The WG 14 of TC 213 is working on ISO/DTS 17450 Geometrical Product Specifications (GPS) - Model for geometrical specification and verification [32, 33]. This is the basis on which the entire GPS language is built. This is the most important and fundamental part. It has been decided to split this standard into two parts

- ISO/CD 17450-1: Geometrical product specifications (GPS)- General concepts- Part-1: Model for geometric specification and verifications;

- ISO/CD 17450-2: Geometrical product specifications (GPS)- General concepts- Part-2: Operators and uncertainties.

In this standard they are classifying the various features used in the field of GPS and defining the operators on these features. For example, ideal feature: perfect shape feature defined by a type (geometrical properties of a feature) and characteristics (parameter of one feature or between two features expressed as a length or an angle), non-ideal feature: imperfect feature fully dependent of the skin model, and skin model: model of the the physical workpiece with its environment. The following operations are defined in ISO/CD 17450 [32, 33].

Partition: used to identify bounded features.

Extraction: used to identify finite number of points from a feature, with specific rules.

Collection: used to consider some features together, which together play a functional role. It is possible to build the collection of ideal features or the collection of non-ideal features. (Composite feature)

Construction: An operation called construction is used to build ideal features from other features maintaining the constraints.

WG 14 is working on "Operator Principle" by which they wish to study whether the set of features along with these operators can be given a good mathematical structure. But they base their assumption that controlling geometry will imply controlling function. Controlling geometry will at the most imply controlling geometric part of the function. Geometry is one aspect and the other important aspect is material. The second major assumption they have is that summation of part function will lead to product function, which is not true (there may be some few exceptions/chance cases). There are lot of mathematical and logical issued still unresolved.

The FAB architecture described in this report can substantially help in resolving these issues. Using the various Class structures like **Function**, and **Artifact** we can clearly define functional aspects of a feature without ambiguity. The work of STEP part ISO 10303-47 [26] needs to be harmonized with the work of ISO TC 213. The FAB architecture can do this very effectively by using these class structures.

6 Suggestions for Future Research

In this report, a conceptual framework for representation of artifacts, function, and behavior have been discussed and also issues related to function to form mapping in the conceptual design stages. However, the conceptual framework presented here needs further investigation and analysis in respect to:

- detailed function-form-behavior specification
- development of computational methodologies
- further considerations in detailed tolerance analysis and synthesis
- design optimization considering various aspects affecting the product life cycle starting from the user specification to manufacturing.

A detailed specification and formal definition of artifacts in terms of its parameters representing functional requirements, shape, goal, constraints, etc, would be developed for putting the concepts discussed in this report on a concrete footing. Representation of functions both as abstract entities and subsequent links to physical objects with concrete shape and behavior, needs more rigorous treatment so that formal qualitative as well as quantitative

reasoning techniques can be used to drive the design process in a systematic manner. The process of selection of sub-functions and/or objects during a functional decomposition and synthesis process can be based on qualitative reasoning using Case-Based Reasoning (CBR) and other methods.

Further work will be required to carry out detailed development of object-oriented tools for the integration of tolerance classes at the conceptual design stage and development of procedures for tolerance synthesis and analysis models.

The various object-oriented models mentioned in this report would be used to investigate the development of an optimal product design schema based on integration of requirements, constraints and goals in all aspects of product life cycle starting from user specification to design, manufacturability, after-sales support and recycling, etc.

Figure 17 describes the overall future directions of this research. The center of it all is the Knowledge Aided Design. This shared knowledge base drives the rest of the applications and tools. This is a clear shift from the present CAD systems which are mostly geometry driven to more information and knowledge-driven design environment. The CAD tools, Analysis/Simulation tools, Process planning tools (CAPP) and NIST developed Process Specification Language (PSL) [40] standard interface, business tools (PDM, ERP, Supply Chain Management) are all connected to the central knowledge base. The various data exchange standards (existing, under development and future needs) are also illustrated in the figure. Apart from these tools the Augmented CAD tools which include virtual reality (VR) tools is also included in the future goal. Various researchers [15, 16, 48] have developed VR tools for assembly/disassembly, motion planning and simulation purposes. The data exchange mechanisms and standards are becoming a major issue. In developing the FAB-DFT model we have provided open connectivity to these standards.

7 Conclusions

A comprehensive information model, FAB model, is described in this report. This model is generic in the sense that it can be used for capturing other information as well. The Unified Modeling Language (UML) [8] model under development will be used as a proof of concept and will eventually lead to standards. The model can also be used to study the various interoperability issues and manufacturing system integration. In [53], a multi-level approach called the Design for Tolerance [DFT] process was proposed which enables tolerancing to be addressed at successive stages of design in an incremental fashion. The effective use of the FAB and DFT model for design tolerancing, starting from conceptual stage of the design and continuously evolving throughout the entire design process to the final detailed design is also discussed in this report.

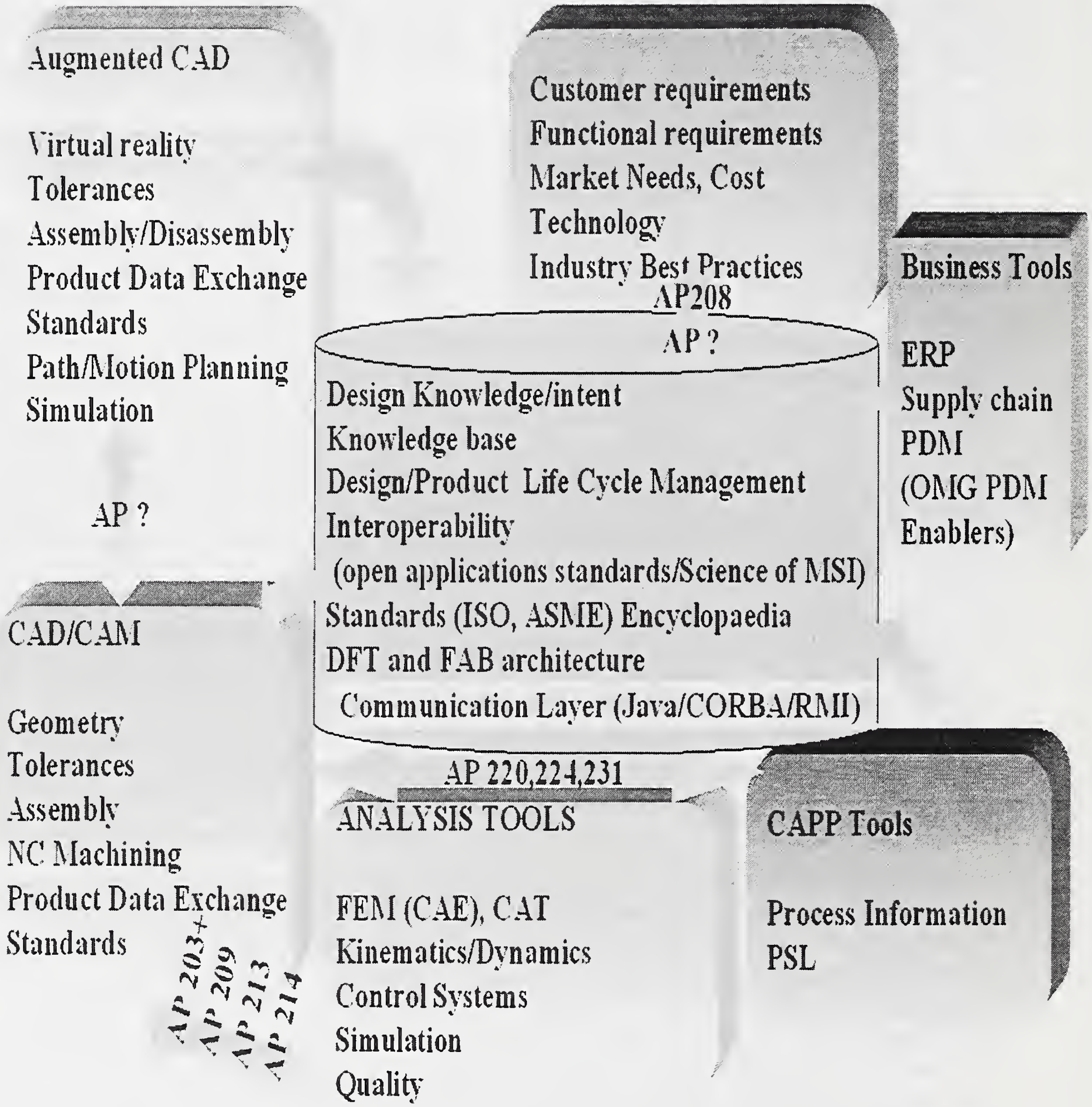


Figure 17: Future research directions - AP: Application Protocol

Acknowledgments

This project is funded [in part] by NIST's Systems Integration for Manufacturing Applications (SIMA) Program. Initiated in 1994 under the federal government's High Perfor-

mance Computing and Communications effort, SIMA is addressing manufacturing systems integration problems through applications of information technologies and development of standards-based solutions. With technical activities in all of NIST's laboratories covering a broad spectrum of engineering and manufacturing domains, SIMA is making information interpretable among systems and people within and across networked enterprises. This work is also sponsored by the RaDEO (Rapid Design Exploration and Optimiztion) program at DARPA (Defense Advanced Research Project Agency).

Disclaimer

No approval or endorsement of any commercial product, services, or company by the National Institute of Standards and Technology is intended or implied.

References

[1] **ASME.** *Dimensioning and Tolerancing (Y14.5M).* ISBN 0791822230. ASME (1994).

[2] **ASME.** *Mathematical Definitions of Dimensioning and Tolerancing Principles (Y14.5.1M).* ISBN 0791822524. ASME (1994).

[3] **Barkmeyer, E. J.** *SIMA (Systems Integration for Manufacturing Applications) reference architecture, Part 1: Activity models.* Technical report, National Institute of Standards and Technology, Internal Report 5939 (1996).

[4] **Barnes, C. J.**, **Dalgleish, G. F.**, **Jared, G. E. M.**, **Swift, K. G.**, and **Tate, S. J.** *Assembly Sequence Structures in Design for Assembly. Proceedings of the 1997 IEEE Int. Sym. on Assembly and Task Planning*, pp. 164–169. Marina del Rey, CA (1997).

[5] **Baxter, J. E.**, **Juster, N. P.**, and **de Pennington, A.** *Verification of Product Design Specifications using a Functional Data Model. Proceedings of 1994 Lancaster International Workshop on Engineering Design, Computer Aided Conceptual Design* (1994).

[6] **Bender, A.** *Benderizing tolerances – A simple probability method of handling tolerances for limit stack-ups. Graphic Science*, pp. 17–23 (1962).

[7] **Bharadwaj, B.** *A Framework for Tolerance Synthesis of Mechanical Components.* Master's thesis, Syracuse University, Syracuse, NY (1995).

[8] **Booch, G.**, **Rumbaugh, J.**, and **Jacobson, I.** *The Unified Modeling Language User Guide.* Addison Wesley (1998).

[9] **Boothroyd, G.** *Product design for manufacturing and assembly. Computer Aided Design*, volume 26(7):pp. 505–520 (1994).

[10] **Boyles, R. A.** *The Taguchi Capability Index. Journal of Quality Technology*, volume 23(1):pp. 17–26 (1991).

[11] **Chan, L. K.**, **Cheng, S. W.**, and **Spiring, F. A.** *A new measure of process capability C_{pm}. Journal of Quality Technology*, volume 20:pp. 162–175 (1988).

[12] **Chandrasekaran, B.**, **Goel, A. K.**, and **Iwasaki, Y.** *Functional Representation of Design Rationale. Computer*, pp. 48–56 (1993).

[13] **Chase, K. W.** and **Parkinson, A.** *A survey of research in the application of tolerance analysis to the design of mechanical assemblies. Research in Engineering Design*, volume 1(3):pp. 23–37 (1991).

[14] **Clements, J.** *Process capability calculation for non-normal calculation. Quality Progress*, volume 22(2):pp. 49–55 (1989).

[15] **Connacher, H.**, **Jayaram, S.**, and **Lyons, K. W.** *Integration of Virtual Assembly with CAD. Symposium on Virtual Reality in Manufacturing Research and Education* . Chicago, IL (1996).

[16] **Connacher, H.**, **Jayaram, S.**, and **Lyons, K. W.** *Virtual Assembly Using Virtual Reality Technique. CAD* (Accepted).

[17] **Evans, D. H.** *Statistical Tolerancing: The State of the art, Part I. Background. Journal of Quality Technology*, volume 6:pp. 188–195 (1974).

[18] **Evans, D. H.** *Statistical Tolerancing: The State of the art, Part II. Method for estimating moments. Journal of Quality Technology*, volume 7:pp. 1–12 (1975).

[19] **Gilson, J.** *A new approach to engineering tolerances.* Machinery Pub. Co. Ltd., London (1951).

[20] **Gorti, S. R.** and **Sriram, R. D.** *From symbol to form: A framework for conceptual design. Computer Aided Design*, volume 28(11):pp. 853–870 (1996).

[21] **Gupta, S. K.**, **Paredis, C. J. J.**, **Sinha, R.**, **Wang, C. H.**, and **Brown, P. F.** *Intelligent Assembly Modeling and Simulation.* ICES 05-104-97, Carnegie Mellon University, Carnegie Mellon University, Pittsburgh, PA (1997).

[22] **Harry, M. J.** *The nature of six sigma quality.* Technical report, Government Electronics Group, Motorola Inc., Scottsdale, Arizona (1987).

[23] **Harry, M. J.** and **Lawson, J. R.** *Six Sigma Producibility Analysis and Process Characterization.* Motorola University Press (1991).

[24] **Harry, M. J.** and **Stewart, R.** *Six sigma mechanical design tolerancing.* Technical report, Government Electronics Group, Motorola Inc., Scottsdale, Arizona (1988).

[25] **Iannuzzi, M. P.** and **Sandgren, E.** *Tolerance optimization using genetic algorithms: Benchmarking with manual analysis. Proceedings of the Fourth CIRP Design Seminar on Computer-Aided Tolerancing*, (Edited by **Kimura, F.**), pp. 220–234. Chapman & Hall, London (1996).

[26] **ISO:10303-47**. *Industrial automation systems and integration- Product data representation and exchange- ISO 10303-47:Integrated generic resources: Shape variation tolerances.* http://www.iso.ch (1997).

[27] **ISO:10360-2**. *Coordinate metrology – Part 2: Performance assessment of coordinate measuring machines.* ISO (1994).

[28] **ISO/CD:5459-1**. *ISO/CD 5459-1: Geometrical Product Specifications (GPS)- Datums for geometrical tolerancing- Part 2: Datums and Datum systems: General Terms and Definitions.* http://www.iso.ch/ (Revision of ISO 5459:1981).

[29] **ISO/CD:5459-2**. *ISO/CD 5459-2: Geometrical Product Specifications (GPS)- Datums for geometrical tolerancing- Part 2: Datums and Datum systems: Explanations and Indications.* http://www.iso.ch/ (Revision of ISO 5459:1981).

[30] **ISO/CD:5459-3**. *ISO/CD 5459-3: Geometrical Product Specifications (GPS)- Datums for geometrical tolerancing- Part 2: Datums and Datum systems: Method for the establishment of datums for the assessment of geometrical tolerances.* . http://www.iso.ch/ (Revision of ISO 5459:1981).

[31] **ISO/DIS:1101**. *ISO/DIS 1101: Geometrical Product Specification (GPS) - Geometrical tolerancing - Generalities, definitions, symbols, indication on drawings. (Revision of ISO 1101:1983).* http://www.iso.ch/ (Under Revision).

[32] **ISO/DTS:17450-Part:1**. *Geometrical Product Specification (GPS)- Model for geometric specification and verifications.* Technical report, http://www.iso.ch/ (Under Revision).

[33] **ISO/DTS:17450-Part:2**. *Geometrical Product Specification (GPS)- Operators and uncertainties* . Technical report, http://www.iso.ch/ (Under Revision).

[34] **ISO/TC/213**. *Dimensional and geometrical product specification and verification, http://www.ds.dk/isotc213* (1999).

[35] **Johnson, A. L.** *Functional Modeling: A New Development in Computer-Aided Design. Intelligent CAD*, (Edited by **Yoshikawa, H.** and **Holden, T.**), volume II, pp. 203–212. Elsevier Science Publishers B.V., North Holland, Amsterdam (1990).

[36] **Kacker, R. N.** *Off-line quality control, parameter design, and the Taguchi method. Journal of Quality and Technology*, volume 17(4):pp. 176–188 (1985).

[37] **Kacker, R. N.** *Taguchi Methods. Handbook of Statistical Methods for Engineers*, pp. 19.1–19.27. Mc-Graw Hill (1988).

[38] **Kanai, S.**, **Onozuka, M.**, and **Takahashi, H.** *Optimal tolerance synthesis by genetic algorithm under the machining and assembling constraints. Proceedings of the Fourth CIRP Design Seminar on Computer-Aided Tolerancing*, (Edited by **Kimura, F.**), pp. 235–250. Chapman & Hall, London (1996).

[39] **Kane, V. E.** *Process capability indices. Journal of Quality and Technology*, volume 18(1):pp. 41–52 (1986).

[40] **Knutilla, A.**, **Schlenoff, C.**, **Ray, S.**, **Polyak, S. T.**, **Tate, A.**, **Cheah, S. C.**, and **Anderson, R. C.** *Process Specification Language: An Analysis of Existing Representations.* Technical Report NISTIR 6160, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, USA (1998).

[41] **Ko, H.** and **Lee, K.** *Automatic Assembling Procedure Generation from Mating Conditions. CAD*, volume 21(9):pp. 561–569 (1989).

[42] **Kotz, S.** and **N.L.Johnson**. *Process Capability Indices.* Chapman & Hall (1993).

[43] **Kusiak, A.** and **Feng, C. X.** *Deterministic tolerance synthesis: A comparative study. Computer Aided Design*, volume 27(10):pp. 759–768 (1995).

[44] **Lai, K.** and **Wilson, W. R. D.** *FDL - A Language for Function Description and Realization in Mechanical Design . Journal of Mechanisms, Transmissions, and Automation in Design, Transactions of the ASME*, volume 111:pp. 117–123 (1989).

[45] **Lee, J.** and **Johnson, G. E.** *Optimal tolerance allotment using a genetic algorithm and truncated Monte Carlo simulation. Computer Aided Design*, volume 25(9):pp. 601–611 (1993).

[46] **Lee, J.**, **Woo, T. C.**, and **Chou, S. Y.** *Tolerance synthesis for non-linear systems based on non-linear programming. IIE Transactions*, volume 25(1):pp. 51–61 (1993).

[47] **Lee, K.** and **Gossard, D. C.** *A hierarchical data structure for representing assemblies: Part 1. Computer Aided Design*, volume 17(1):pp. 15–19 (1985).

[48] **Lyons, K.**, **Shooter, S.**, **Keirouz, W.**, and **Hart, P.** *The Open Assembly Design Environment: An Architecture for Design Agent Interoperability. Proceedings of DETC 99, Paper No. DFM-8945.* ASME, Las Vegas, Nevada, USA (1999).

[49] **Mantripragada, R.**, **Cunningham, T. W.**, and **Whitney, D. E.** *Assembly oriented design: A new approach to designing assemblies. Product Modeling for Computer Integrated Design and Manufacture*, (Edited by **Pratt, M. J.**, **Sriram, R. D.**, and **Wozny, M. J.**), pp. 308–324. Chapman & Hall (1997).

[50] **M.Johnson**. *Statistics Simplified. Quality Progress*, volume 25(1):pp. 10–11 (1992).

[51] **M.J.Pratt.** *Parameterization and constraints for explicit geometric product models.* Technical Report N184, ISO TC184/SC4/WG12 (1998). (http://www.nist.gov/sc4/wg_qc/wg12/n184/).

[52] **Murdock, J. W.**, **Szykman, S.**, and **Sriram, R. D.** *An Information Modeling Framework to Support Design Databases and Repositories. Proceedings of DEC97. 1997 ASME Design Engineering Technical Conferences, Sacramento, California* (1997).

[53] **Narahari, Y.**, **Sudarsan, R.**, **Lyons, K. W.**, **Duffey, M. R.**, and **Sriram, R. D.** *Design for Tolerance of Electro-mechanical Assemblies: An Integrated Approach.* NISTIR 6223, National Institute of Standards and Technology, NIST, Gaithersburg, MD 20899 (1998).

[54] **Narahari, Y.**, **Sudarsan, R.**, **Lyons, K. W.**, **Duffey, M. R.**, and **Sriram, R. D.** *Design for Tolerance of Electro-mechanical Assemblies: An Integrated Approach. IEEE Transactions on Robotics and Automation*, volume 15(6):pp. 1062–1079 (1999).

[55] **Parks, J. M.** *The Unified Methodology on Design for latitude as a Concept and Capability for Tolerance Analysis and Synthesis. Selected Case Studies in the Use of Tolerance and Deviation Information during Design of Representative Industrial Products*, (Edited by **Tipnis, V.**), pp. 129–150. ASME, New York (1992).

[56] **Parks, J. M.** *The two enablers for Holistic Probabilistic Design: Methodology and a complete suite of tools for stochastic analysis. Proceedings of the Third International Conference on Stochastic Structural Dynamics* (1995).

[57] **Parks, J. M.** *Holistic Approach and Advanced Techniques and Tools for Tolerance Analysis and Synthesis. Computer-Aided Tolerancing*, (Edited by **Kimura, F.**), pp. 251–269. Chapman and Hall, London (1996).

[58] **Parks, J. M.** *Advanced Methodology and Software for Tolerancing and Stochastic Optimization. Proceedings of the Fifth International Seminar on Computer-Aided Tolerancing, Toronto, Canada* (1997).

[59] **Phadke, M. S.** *Quality Engineering using Robust Design.* Prentice Hall, Englewood Cliffs, NJ, USA (1989).

[60] **Qu, G.** *Automatic Tolerance Allocation with Process Selection Based on a Hierarchical Assembly Model.* Ph.D. thesis, University of Minnesota, University of Minnesota, Minneapolis (1993).

[61] **Rajan, V. N.**, **Lyons, K. W.**, and **Sreerangam, R.** *Identification of requirements and representation for assembly modeling. Proceedings of the 1997 IEEE International Symposium on Assembly and Task Planning.* Marina del Ray, CA (1997).

[62] **Roy, U.**, **Banerjee, P.**, and **Liu, C. R.** *Design of an automated assembly environment. CAD*, volume 21(9):pp. 561–569 (1989).

[63] **Roy, U.** and **Bharadwaj, B.** *Tolerance Synthesis in a Product Design System.* Technical Paper MS96-146. North American Manufacturing Research Institution. Society of Manufacturing Engineers, Dearborn, MI (1996).

[64] **Roy, U.**, **Bharadwaj, B.**, **Chavan, A.**, and **Mohan, C. K.** *Development of a Feature Based Expert Manufacturing Process Planner. Proc. of the 7th IEEE International Conference on Tools with Artificial Intelligence*, pp. 65–73 (1995).

[65] **Roy, U.** and **Li, B.** *Representation and interpretation of Geometric tolerances for polyhedral objects, Part I:Form Tolerances. CAD*, volume 30(2):pp. 151–161 (1998).

[66] **Roy, U.** and **Li, B.** *Representation and interpretation of Geometric tolerances for polyhedral objects, Part II:Size, Orientation, and Positional Tolerances. Submitted for Publication* (1998).

[67] **Roy, U.** and **Liu, C. R.** *Establishment of functional relationships between product components in assembly database. CAD*, volume 20(10):pp. 570–580 (1988).

[68] **Roy, U.**, **Sudarsan, R.**, **Sriram, R. D.**, **Lyons, K. W.**, and **Duffey, M. R.** *Information Architecture for Design Tolerancing: From Conceptual to the Detail Design. Proceedings of DETC 99, Paper No. DAC-8704.* ASME, Las Vegas, Nevada, USA (1999).

[69] **Spiring, F. A.** *A unifying approach to process capability indices. Journal of Quality Technology*, volume 29(1):pp. 49–58 (1997).

[70] **Srikanth, S.** and **Turner, J. U.** *Toward a unified representation of mechanical assembly. Engineering with Computers*, volume 6(2):pp. 103–112 (1990).

[71] **Srinivasan, V.** *A geometrical product specification language based on a classification of symmetry groups.* Technical Report ISO/TC 213/WG 14 N 26, IBM Watson Research Center (1998).

[72] **Srinivasan, V.** and **O'Connor, M. A.** *On interpreting statistical tolerancing. Manufacturing Review*, volume 7(4):pp. 304–311 (1994).

[73] **Srinivasan, V.** and **O'Connor, M. A.** *Towards an ISO standard for statistical tolerancing. Proceedings of the Fourth CIRP Design Seminar on Computer-Aided Tolerancing*, (Edited by **Kimura, F.**), pp. 159–172. Chapman & Hall, London (1996).

[74] **Srinivasan, V.**, **O'Connor, M. A.**, and **Scholz, F. W.** *Techniques for composing a class of statistical tolerance zones. Advanced Tolerancing Techniques*, (Edited by **Zhang, H.-C.**), Wiley Series in Engineering Design and Automation, pp. 139–166. John Wiley & Sons (1997).

[75] **Srinivsan, V.** *Statistical Specification of Geometrical characteristics. KC Symposium 98.* Kennedy Space Center, FL (1998).

[76] **Sudarsan, R.**, **Lyons, K. W.**, and **Pratt, M. J.** (Editors). *Assembly Level Tolerancing: Standards and Implementation Issues.* Workshop Handout, National Institute of Standards and Technology (1997).

[77] **Sugimura, N.** *A study on product model for design and analysis of mechanical assemblies based on STEP.* Proceedings of the Japan-USA symposium on Flexible Automation (1994). Kobe-Japan.

[78] **Taguchi, G.** *Introduction to Quality Engineering.* Asian Productivity Organization (1986).

[79] **Taguchi, G.** *Taguchi on Robust Technology.* ASME Press, New York (1993).

[80] **Tsai, J.-C.** and **Cutkosky, M. R.** *Representation and reasoning of geometric tolerances in design. AI-EDAM*, volume 11:pp. 325–341 (1997).

[81] **Wang, N.** and **Ozsoy, T. M.** *Representation of assemblies for automatic tolerance chain generation. Engineering with Computers*, volume 6(2):pp. 121–126 (1990).

[82] **Whitney, D. E.**, **Gilbert, O. L.**, and **Jastrzebski, M.** *Representation of geometric variations using matrix transforms for statistical tolerance analysis in assemblies. Research in Engineering Design*, volume 6:pp. 191–210 (1994).

Appendix A

A Process Capability Indices (PCIs): A brief tutorial

A manufacturing process is characterized by the measurements of process operations and units output from the process. In most cases these measurements are assumed approximately normal so that a process under statistical control may be understood by first two moments mean μ and standard deviation σ of the statistical distribution.

A process specification usually consists of lower and upper specification limits (LSL, USL) and a target T somewhere between the limits (most often $T = MSL = \frac{1}{2}(LSL+USL)$ which is a symmetric case). A capability index (random variable) is a unitless function of the process parameters (μ, σ) and the process specification (LSL, T, USL) designed as *a tool to aid in the assessment of process performance* [39]. Computation of the capability index involves finding the ratio of allowable process spread or tolerance and actual process spread or variation. For any particular process the allowable process spread is a fixed requirement, while the actual process spread is generally an unknown quantity that will have to estimated. The two main approaches for PCIs are:

1. based on expected proportion of non-conforming items (C_p and C_{pk});
2. based primarily on loss function (C_{pm}).

The first approach necessitates the assumption on the form of the distribution while the second approach assumes the form of loss function. In this report our discussion on PCIs is limited to the basic definitions, properties, and usage as a quality tool. For a complete discussion about the statistical properties and estimation procedures, interested reader can refer to [42] and the references cited in that monograph.

A.1 The Basic PCIs: C_p and C_{pk}

Suppose the lower and upper specification limits for the value of a measured characteristic X be specified as LSL and USL respectively. Values of X outside $[LSL, USL]$ will be termed as non-conforming and we are interested in the probability

$$Pr\{LSL \leq X \leq USL\}.$$

An indirect measure of potential ability (capability) to meet the requirement $(LSL \leq X \leq USL)$ is the process capability index C_p:

$$C_p = \begin{cases} \frac{USL-LSL}{6\sigma} & \text{Bilateral} \\ \frac{USL-\mu}{3\sigma} & \text{Unilateral with } USL \text{ known} \\ \frac{\mu-LSL}{3\sigma} & \text{Unilateral with } LSL \text{ known} \end{cases} \tag{7}$$

where σ denotes the standard deviation of X. From (7) we can clearly see the large values of C_p are desirable. From the definition of C_p we see

$$\text{Capability index} = \frac{\text{Specification width}}{\text{Distribution width}} = \frac{\text{allowable process spread or tolerance}}{\text{actual process spread or variation}}.$$

Usually it is assumed that manufacturing process measurements follow normal distribution, that is, $X \sim N(\mu, \sigma)$ and often it is assumed that target is the mid-point of the specification limits, that is, $T = MSL = \frac{1}{2}(LSL + USL)$ which is a symmetric case.

The use of factor 6 in equation (7) has the following implication: if $C_p = 1$ and the process distribution is normal with $\mu = \frac{1}{2}(LSL + USL)$ then the proportion of non-conforming products is as small as 0.27% which is equivalent to 2700 ppm (parts per million) and this has been a standard practice in statistical process control (SPC).

If the distribution is assumed to be normal and if $\mu = T = \frac{1}{2}(LSL + USL)$, then the expected proportion of non-conforming products is given by $2\Phi(-\frac{d}{3\sigma})$, where $\Phi(.)$ is the standard normal CDF (cumulative density function) and $d = \frac{(USL-LSL)}{2}$ is the half-length of the specification width. From this we can see that the percentage yield is given by $100[2\Phi(3C_p) - 1]$. If $C_p = 1$ the expected proportion of non-conforming products is 0.27%, which is regarded as acceptably small. For acceptance we should have $C_p \leq c$ with $c = 1, 1.33, 1.5, 2.0$ corresponding to $(USL - LSL) = 6\sigma, 8\sigma, 9\sigma, 12\sigma$. It is important to note that $C_p = 1$ does not guarantee that there will be only 0.27% of non-conforming product. In fact, all that it does guarantee is that, with assumption of normality and the relevant value of σ, there will never be less than 0.27% expected proportion of non-conforming product. It is only when $\mu = T = \frac{1}{2}(LSL + USL)$ that the expected proportion of non-conforming product is small as 0.27%. A good discussion of these indices including estimating procedures, sampling strategies, and application methodologies is given in [39].

A.2 C_{pk} Index

It is clear from the definition of C_p, it fails to take into consideration the effect of $|\mu - T|$. Because of inability of C_p to consider target, several indices have been proposed that take into account the target T. The C_{pk} index was introduced to give the values of μ some

influence on the value of the PCI. C_{pk} is defined as

$$\begin{aligned} C_{pk} &= \frac{\min\{USL-\mu, \mu-LSL\}}{3\sigma} \\ &= \frac{d-\left|\mu-\frac{(USL+LSL)}{2}\right|}{3\sigma} \quad (\text{since} \quad \min(a,b) = \tfrac{1}{2}(|a+b|-|a-b|) \\ &= \left\{1-\frac{\left|\mu-\frac{(USL+LSL)}{2}\right|}{d}\right\} C_p \end{aligned} \tag{8}$$

Since $C_p = \frac{d}{3\sigma}$, we have $C_{pk} \leq C_p$, with equality *if and only if* $\mu = T = \frac{1}{2}(LSL+USL) = MSL$. It seems that the original way of compensating for the lack of input μ in the calculation of C_p, was to use an additional statistic k (Centering Index) defined as

$$\begin{aligned} k &= \frac{|\mu-MSL|}{d} \\ &= \frac{|\mu-T|}{d} \quad \text{for any general Target } T \end{aligned} \tag{9}$$

This is equivalent to using μ (through k) and σ(through C_p) separately. While it might be said that one might as well use μ and σ themselves (or their estimates), k and C_p do have the advantage of being dimensionless, and related to the specification limits. In [72], the centering index k is denoted as C_c. Johnson [50] even points out that "none of the indices adds any knowledge or understanding **beyond** that contained in the basic parameters μ, σ, T, LSL, USL.

With the combination into the index $C^*_{pk} = (1-k)C_p$, the reduction to single index indeed had to be paid for by losing separate information on location (μ) and dispersion (σ). C_{pk} and C^*_{pk} are numerically equivalent if $0 \leq k \leq 1$. C_{pk} captures the off-targetness. The centering index k represents the amount of tolerance consumed by the mean shift.

Yet another way of defining C_{pk} is by defining C_{pl} and C_{pu} as

$$C_{pl} = \frac{\mu - LSL}{3\sigma}$$

and

$$C_{pu} = \frac{USL - \mu}{3\sigma}$$

and define

$$C_{pk} = \min\{C_{pl}, C_{pu}\}$$

If the distribution is assumed to be normal, then the expected proportion of non-conforming product is

$$\Phi\left(\frac{\mu - LSL}{\sigma}\right) + \left\{1 - \Phi\left(\frac{USL-\mu}{\sigma}\right)\right\}$$

and we have

$$100[\Phi(3C_{pk}) - 1] \leq \%\text{Yield} \leq 100[\Phi(3C_{pk})].$$

A.3 C_{pm} Index

The C_p and C_{pk} indices are appropriate measures of progress for quality improvement in which 'reduction of variability' is the main criterion and process yield is the primary measure of success. Taguchi [78, 79] has suggested a different approach to quality improvement in which 'reduction of variation from the target value' is the guiding principle. The index C_{pm} is defined as

$$\begin{aligned} C_{pm} &= \frac{USL-LSL}{6\sqrt{\sigma^2+(\mu-T)^2}} \\ &= \frac{d}{3\sqrt{\sigma^2+(\mu-T)^2}} \end{aligned} \tag{10}$$

The cost of a characteristic X missing the target T is often assumed to be well approximated by the symmetric squared error loss function

$$loss(X) = c(X-T)^2$$

where c is some positive constant. Since $E[(X-T)^2] = E(X^2) + T^2 - 2\mu T + \mu^2 - \mu^2 = \sigma^2 + (\mu - T)^2 = E[loss(X)]$, the C_{pm} index can be written as

$$C_{pm} = \frac{USL - LSL}{6\sqrt{E[loss(X)]}}.$$

The relationship of C_{pm} to squared error loss is discussed in [50]. Chan *et al* [11] proposed the index C_{pm} with the assumption $\mu = T$. Boyles [10] extended the definition without the restrictive assumption $\mu = T$ and explains how a given value of C_{pm} places an upper-bound on $|\mu - T|$.

Since $\sigma^2 + (\mu - T)^2 \geq \sigma^2$ we have $C_{pm} \leq C_p$ and $C_{pm} = C_p \iff \mu = T$. The index C_{pm} as defined is not a satisfactory measure of process capability unless the target value is equal to the mid-point of the specification limit [42]. A modified index C^*_{pm} is defined to overcome this problem.

$$C^*_{pm} = \frac{\min\{USL - T, T - LSL\}}{3\sqrt{\sigma^2 + (\mu - T)^2}}$$

If $T = MSL$ then $C_{pm} = C^*_{pm}$.

A.4 Comparison of PCIs

If $\mu = T$, then $C_p = C_{pk} = C_{pm}$. If the process distribution $\sim N(\mu, \sigma)$ (and in statistical control) then a value of 1 for all the PCIs means that the expected proportion of product within the specification limits is $\geq 99.7\%$. Negative values are not permitted for PCIs.

As already explained, the main function of C_{pk} is to indicate the degree to which the process is within specification limits. For $LSL < \mu < USL$, $C_{pk} \to \infty$ as $\sigma \to 0$, but large

values of C_{pm} do not provide information about $|\mu - T|$. The index $C_{pm} \to \infty$ if $\mu \to T$ as well as $\sigma \to 0$ and $C_{pm} \to 0$ as $|\mu - T| \to \infty$. The index C_{pm} is bounded above as

$$C_{pm} < \frac{USL - LSL}{6|\mu - T|}. \tag{11}$$

It is clear from (11), when $T = MSL$, a C_{pm} value of 1 implies that the process mean μ lies within the middle third of the specification width.

In [69] a unifying approach is discussed that ties the various indices together. Through the use of a weight function the relationship that exists among C_p, C_{pk}, C_{pm} and C^*_{pk} is discussed [69]. The index C_{pw} is defined as

$$C_{pw} = \frac{USL - LSL}{6\sqrt{\sigma^2 + w(\mu - T)^2}} \tag{12}$$

where w represents the weight function, $C_{p0} = C_p$ and $C_{p1} = C_{pm}$. For complete details refer [69]. There are other additional weight functions, for example,

$$w = \begin{cases} r & 0 < r \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$$w = \begin{cases} \frac{1}{r} & 0 < r \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

The weight function (13) would typically be used in those situations where minor departures from the target are not of interest, but major departures are considered critical. The weight function (14) applies greater weight to minor departures from the target, while only a marginal decrease in capability results once the process deviates from the target by more than 1σ. Using different weight functions allows one to customize the PCI to the process of interest, thereby allowing different shaped loss functions (as in the case of C_{pm}).

A.5 Example

Suppose we have three different lots of fasteners (not measured) and we are interested in their values (say, length). We can compute the sample mean $\bar{x} = \frac{1}{n}\sum x_i$, where n is the sample size. We assume, for the sake of simplifying our discussion, that sample mean and variance is equal to population mean and variance. This implies that our sample is a perfect representation of the population, which we assume as a Gaussian population.

Let the specification limits be specified as $LSL = 114$ mm, $USL = 126$ mm and the target $T = 120$ mm. Let $\mu_1 = 120, \mu_2 = 123, \mu_3 = 124$ and $\sigma_1 = 2, \sigma_2 = 1, \sigma_3 = 0.6666$ be the means and standard deviations of three populations respectively. From the theory

of the normal distribution, we can compute the following probabilities (Lot#1 is taken as example).

$\mu \pm 1\sigma$	68.26%	Pr{fastener $\in$ [118,122]} = 0.6826
$\mu \pm 2\sigma$	95.46%	Pr{fastener $\in$ [116,124]} = 0.9546
$\mu \pm 3\sigma$	99.73%	Pr{fastener $\in$ [114,126]} = 0.9973
$\mu \pm 6\sigma$	99.9999998%	Pr{fastener $\in$ [108,132]} = .999999998

From the table we can see that 99.73% of the fasteners lie in [114,126] mm. The expected proportion of non-conforming products is 0.27% or the defect rate is 2700 ppm (parts per million). [2700 ppm is computed as $(1 - 0.9973) \times 1000000$]. The defect rate for Lot#1 would be 0.002 ppm if we have $\sigma_1 = 1$. Let us now compute the PCIs C_p, C_{pk} and C_{pm}.

	Lot#1	*Lot#2*	*Lot#3*
k	0	$\frac{\lvert 123-120 \rvert}{6} = 0.5$	$\frac{\lvert 124-120 \rvert}{6} = \frac{2}{3}$
C_p	1	2	3
C_{pk}	1	1	1
$q = \frac{\lvert \mu - T \rvert}{\sigma}$	0	3	6
$C_{pm} = \frac{1}{\sqrt{1+q^2}}$	1	0.63	0.49

Note that C_{pm} clearly distinguishes the three different lots. The index C_p is increasing from Lot#1 to Lot#3 while C_{pm} is decreasing from Lot#1 to Lot#3. It is important to note that the mean is drifting from the target for lots two and three. The index C_{pk} is equal to one for all the lots and does not capture process centering.

A.6 Generalization of PCIs

The previous discussion about PCIs assume the normal distribution for the process (or for the measured characteristic X), which is a strong assumption. We would like to drop this normality assumption. This may be necessary, because perfect normality is seldom realized in practice, and even in theory the normal distribution might not be expected for certain situations (for example, perpendicularity tolerance).

The study of non-normal distribution for X falls into two main parts. The first is the investigation of the properties of PCIs and their statistical estimates when the distribution has specific non-normal forms. The second, more difficult, is development of methods of allowing for non-normality and consideration of use of new PCIs specifically designed to be robust (that is, not too sensitive) to non-normality. There are several methods suggested in the literature, here we discuss briefly about (1) Clements' Method and (2) Johnson-Kotz-Pearn Method. For complete details refer [42].

A.6.1 Clements' Method

Clements [14] proposed a method of construction based on the assumption that the process distribution can adequately be represented by Pearson distribution. His suggestion was to replace the multiplier '6' in the denominator of C_p by number θ such that

$$Pr\left\{\mu - \frac{1}{2}\theta\sigma \leq X \leq \mu + \frac{1}{2}\theta\sigma\right\} = 0.0027.$$

For given values of skewness ($\sqrt{\beta_1}$) and kurtosis (β_2) coefficients, compute θ such that the above probability is true. For example, if $\sqrt{\beta_1} = 1$ and $\beta_2 = 5$, we have $\theta = 6.572$ [14]. The index C_p would be calculated as $\frac{d}{3.286\sigma}$. For normal distribution we have $\theta = 6$ so that $C_p = \frac{d}{3\sigma}$. For complete discussion about this methods refer [14].

A.6.2 Johnson-Kotz-Pearn (JKP) Method

Clements' method requires the evaluation (estimation) of $\sqrt{\beta_1}$ and β_2 which may be difficult to obtain. We need rather large samples for accurate estimation these quantities. This method defines a new PCI $C_{p(\theta)}$ as [42]

$$C_{p(\theta)} = \frac{USL - LSL}{\theta\sigma} = \frac{d}{\left(\frac{1}{2}\theta\sigma\right)}$$

where θ is chosen so that the 'capability' - namely the proportion of conforming items, with optimum choice of μ is not greatly affected by the shape of the distribution. Note that $C_p = C_{p(\theta=6)}$. A value of 5.15 is recommended for θ [42], so that $C_p = C_{p(\theta=5.15)} = \frac{d}{2.575\sigma}$ and $C_{pk} = \frac{d-|\mu-MSL|}{2.575\sigma}$.

Note that robustness is attained at the cost of abandoning the 0.27% level for expected proportion of non-conforming items (2700 ppm). Both methods rely on the assumption that the population distribution is uni-modal shape close to a Pearson distribution for Clements method, and more restrictively, close to a Gamma distribution for JKP method.

Remark A.1 *The natural variation (shifts and drifts, for example, due to tool wear, temperature etc) can cause the mean to deviate. Motorola suggests the mean shift could be of the order of 1.5σ units from its target [24]. For a product to be Motorola six sigma quality, that is $d = 6\sigma$, it has to have $C_p \geq 2, C_{pk} \geq 1.5$ with a defect rate of 3.4ppm. Section 2.3.1 (page: 19) discuss this in detail.*

The Motorola convention is to use a one-sided mean shift of 1.5σ. The one-sided mean shift is perhaps motivated by common physical phenomena such as tool wear. A shift of 1.5σ is motivated by earlier work by Bender [6]. Also, it is assumed that the process standard deviation is invariant (Appendix B,page: 76).

If $C_p = 2$ and $C_{pk} = 1.5$ (mean shift consumes 25 percent of the tolerance range), the probability of conformance can be shown to be 0.9999966, which is equivalent to 3.4 ppm. Thus $C_p \geq 2$ and $C_{pk} \geq 1.5$ imply six sigma quality, assuming a 1.5σ one sided mean shift.

The table A.6.2 summarize the various capability indices we have discussed.

Remark A.2 *The ISO effort on standards for statistical tolerancing [72, 73] describes the acceptable statistical distributions of the population of features using C_p, C_{pk} and C_{pm}. At a conceptual level statistical tolerancing simply defines a class of acceptable distributions of the population of features but for practical applications the methods to define and compute these classes should be provided. In [73] three different methods of defining the classes are explained: (1) characterizing the distributions by their first two moments; (2) $C_p - C_{pk}$ plot; and (3) CDF based. Refer Section 2.5, page:25.*

Remark A.3 *All process capability indices of this from are determined under the assumption that (1) the process is in statistical control, (2) the target is the mid-point of the specification limits, and (3) the process measurements are normally distributed. If the process is centered and the target is the mid-point of the specification limits, then the indices are equivalent. If the last two assumptions are relaxed (that is non-symmetric and non-normal) care should be taken in interpreting the values of the indices. But as long as the distributions are determined by their first two moments these indices will serve as an useful tool to assess the process performance.*

LSL, USL	Lower and Upper Specification Limit	T	Target specification
MSL	$MSL = \frac{USL+LSL}{2}$	d	$d = \frac{USL-LSL}{2}$
(μ, σ)	Mean and standard deviation of the process distribution		

Index	Definition	Properties
C_p	$\begin{cases} \frac{USL-LSL}{6\sigma} & \text{Bilateral} \\ \frac{USL-\mu}{3\sigma} & \text{Unilateral } USL \text{ known} \\ \frac{\mu-LSL}{3\sigma} & \text{Unilateral } LSL \text{ known} \end{cases}$	$X \sim N(\mu.\sigma),\ \mu = T = \frac{1}{2}(LSL + USL)$ $C_p = 1 \Rightarrow$ expected proportion of non $-$ conforming products $\leq 0.27\%$ Percentage Yield $= 100[2\Phi(3C_p) - 1]$ $C_p \geq \max\{C_{pk}, C_{pm}\}$
C_{pk}	$\min \frac{\{USL-\mu, \mu-LSL\}}{3\sigma}$ $= \frac{d - \lvert\mu - \frac{(USL+LSL)}{2}\rvert}{3\sigma}$ $= \left\{1 - \frac{\lvert\mu - \frac{(USL+LSL)}{2}\rvert}{d}\right\} C_p$	$C_{pk} \leq C_p$ and $C_{pk} = C_p \iff \mu = T = MSL$ $C_{pk} = 0$ if RHS is $-$ ve If$X \sim N(\mu, \sigma), \mu = T = MSL$ $100[2\Phi(3C_{pk}) - 1] \leq \%\text{Yield} \leq 100[\Phi(3C_{pk})]$
k	$k = \frac{\lvert\mu - MSL\rvert}{d}$ $= \frac{\lvert\mu - T\rvert}{d}$ for any general Target T	Centering Index
C^*_{pk}	$(1-k)C_p$	$0 \leq k \leq 1 \Rightarrow C_{pk} = C^*_{pk}$
C_{pm}	$\frac{USL-LSL}{6\sqrt{\sigma^2+(\mu-T)^2}}$ $= \frac{d}{3\sqrt{\sigma^2+(\mu-T)^2}}$ $= \frac{USL-LSL}{6\sqrt{E[loss(X)]}}$	$C_{pm} \leq C_p$ $C_{pm} = C_p \iff \mu = T = MSL$ C_{pm} as defined is not a statistical measure of process capability unless $T = MSL$

Table 2: PCIs Summary

Appendix B

B The Defect Rate of 3.4 ppm and Motorola Convention

A part or an item is classified as defective (non-conforming) if the desired measurement (quality characteristic), denoted by X, is outside the customer's upper specification limit (USL) or lower specification limit (LSL). In addition to specifying the USL and LSL, a customer would also specify a target value, T, which is typically the midpoint of specification width. For example, let us say that the customer's specification on the length of fasteners is 120 ± 12 mm ($USL = 132, LSL = 108, T = 120$ and the half-tolerance width is 12mm). The process that produce these fasteners is such that the lengths are normally distributed. We assume, for the sake of simplifying our discussion, that sample mean and variance is equal to population mean and variance. This implies that our sample is perfect representation of the population, which we assumed it as normal population.

Assuming that the process is centered, the number of defects per million for the different quality levels is given in Table 3.

Quality Level (σ)	Defect Rate (ppm)
1	317310
2	45500
3	2700
3.5	465
4	63
4.25	22
4.5	6.8
4.75	2.0
5	0.6004
5.25	0.1636
5.5	0.0424
5.75	0.01524
6	0.002

Table 3: Quality levels and defect rates

As table 3 shows, there are 2700 defects per million in the 3σ quality level and only 63 defects per million parts in the 4σ quality level and only 2 defects per billion parts in the 6σ quality level. But then how do we get the defect rate of 3.4 ppm for 6σ quality level? We

will now show the actual computation for the 6σ quality level. The figures in the above table are arrived at using the normal probability tables under the assumption that the process is centered. Refer figure 18 to understand the typical areas under the normal curve. The area under the pobabilty curve for any general distribution is given by Tchebychev's inequality. If the process mean is equal to $T = \frac{1}{2}(USL + LSL)$, the process is centered; otherwise it is

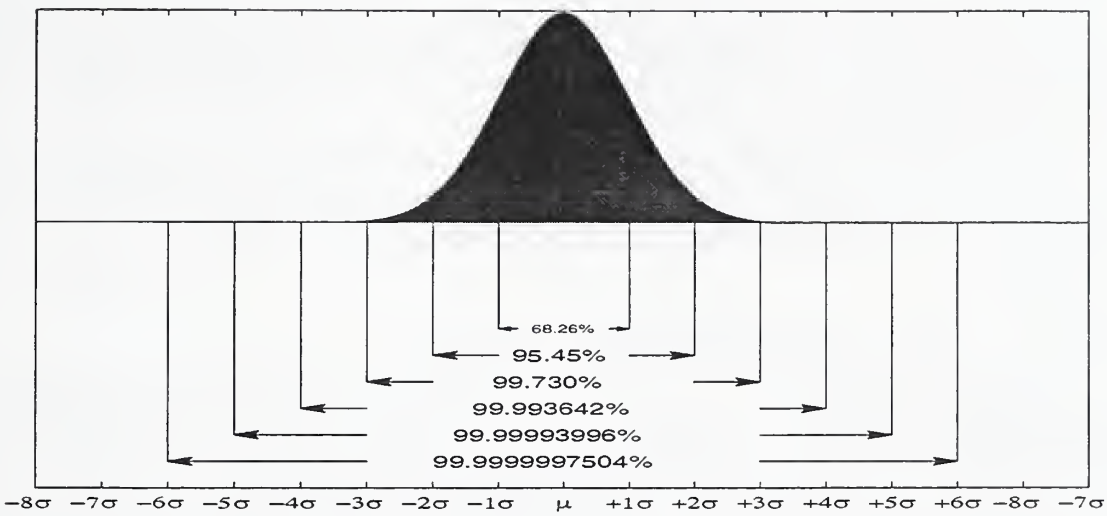


Figure 18: Areas under the normal curve

off-centered. Because of typical shifts and drifts in the process, the process is off-centered and the process mean shifts by as much as 1.5σ units [22]. A shift of 1.5σ is motivated by earlier work by Bender [6] and Gilson [19]. More about this later.

The quality level can be expressed as $m\sigma$, where $m = \frac{(USL-LSL)}{2\sigma}$. Thus if $\sigma = 4$ ($C_p = 1$), the process has 3σ quality level; if $\sigma = 3$ ($C_p = 1.33$), the process has 4σ quality level; and if $\sigma = 2$ ($C_p = 2$), the process has 6σ quality level.

Motorola has established a convention of distributing the total proportion of the non-conformity to the right side of the normal curve. This one-sided mean shift is perhaps motivated by common physical phenomena such as tool wear taht becomes increasingly worse. Also, it is assumed that the process standard deviation is invariant.

The probability calculations with respect to actual are (Figure 19)

$$\begin{aligned} & 1 - Pr\{\text{conformance}\} \\ &= 1 - Pr\{X_A \in [LSL, USL]\} \\ &= 1 - Pr\{108 \leq X_A \leq 132\} \\ &= 1 - Pr\{123 - 15 \leq X_a \leq 123 + 9\} \end{aligned}$$

$$
\begin{aligned}
&= 1 - Pr\{123 - 7.5\sigma \leq X_A \leq 123 + 4.5\sigma\} \\
&= 1 - Pr\{-\infty \leq X_A \leq \mu + 4.5\sigma\} \\
&= 1 - 0.9999966 \\
&= 3.4ppm \text{ defect rate}
\end{aligned}
$$

As said earlier, if $C_p = 2$ and $C_{pk} = 1.5$ (mean shift consumes 25 percent of the tolerance range), the probability of conformance can be shown to be 0.9999966, which is equivalent to 3.4 ppm defect rate. Thus $C_p \geq 2$ and $C_{pk} \geq 1.5$ imply six sigma quality, assuming a 1.5σ one sided mean shift.

In the example, this means that, if the process producing the fasteners has a six-sigma quality level ($\sigma = 2$) and if the process mean is off-centered by as much as 1.5σ units ($117 \leq \mu \leq 123$), the maximum number of defects is 3.4 per million.

B.1 The 1.5σ Mean Shift

As explained in the previous section the six-sigma quality implicitly assumes that the process mean can be off-centered by as much as 1.5σ. It is very difficult to accept that any process would drift as much as 1.5σ without being noticed or corrected.

According to Bender [6], the 1.5σ mean shift is based on "probability, approximation and experience" considerations and sometimes referred as Benderizing. Unfortunately, Bender and Gilson introduced the 1.5σ factor in the different context of tolerancing assemblies of components. Evans [17], in his series of articles, discusses this concept with a classic example of stack of n nominally identical disks. For illustrative purposes, consider $n = 10$ and the height of the stack as the response (assembly response function $Y = X_1 + X_2 + \cdots + X_n$ [53]). Let us assume that the stack height must be 1.25 ± 0.01 in. The problem with statistically tolerancing the components of an assembly (individual disks) deals with setting limits on the allowed excursion of the components' mean and range (using the standard deviation). If σ_i^2 is the variance of the component i, the variance of assembly is given by

$$\sigma_a^2 = Var(\text{assembly}) = \sum \sigma_i^2 = n\sigma_c^2 \quad \text{if } \sigma_i = \sigma_c;\ \forall i \tag{15}$$

Traditionally, the half width of the specified tolerance is set to three times the standard deviation of the assembly in determining the standard deviation of each component. In this example:

$$0.01 = 3\sigma_a = 3\sqrt{10}\sigma_c$$

Thus, the process would be designed so that:

$$\sigma_c = \frac{0.01}{3\sqrt{10}}$$

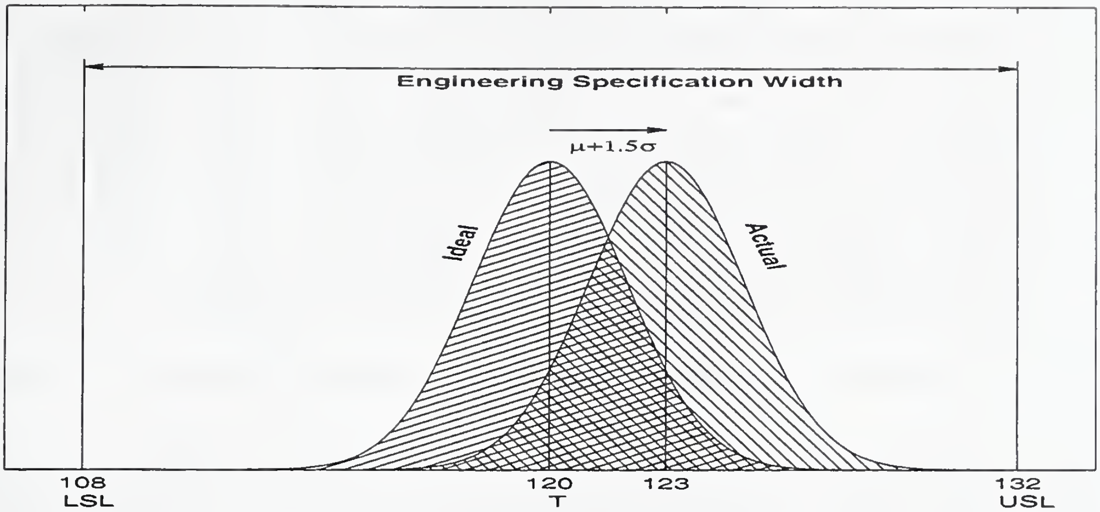


Figure 19: Distribution with 1.5σ mean shift

Bender suggests, however, that due to potential shifts in the means of the disks, it is practical to use $1.5\sigma_a$ as the effective standard deviation of the assembly. Thus, he recommends using $\sigma_a = 1.5\sum \sigma_i^2$ in (1). Harry [22] seems to interpret the factor of 1.5 as the allowance for the shifts in the mean of a single component, due to its being manufactured in different lots. As suggested by Bender and Gilson, it might be reasonable to inflate the estimator of the standard deviation of the assembly to allow for the shifts and drifts in the mean of the individual components. However, we must justify the 1.5σ shift in the process mean of the individual components.

In many situations, adjusting the process to move the process mean closer to the target value is relatively easier and cost-effective than improving the process to reduce the variance. Thus, if the goal is to reduce the number of defects, it does not make sense to improve the process to six-sigma levels and not center the process.

It is interesting to note that the desired quality levels (expressed by the number of defects in ppm) might be achieved through several combinations of off-centering and process standard deviation. For example, a quality level of 3.4 defects per million parts can be achieved in atleast three different ways: 0.5σ off-centering with 5σ quality; 1σ off-centering with 5.5σ quality; and 1.5σ off-centering with 6σ quality. How to achieve a specified quality level or a given number of defects per million depends on the costs associated with adjusting the process mean versus reducing the process variance. If the process mean can be centered, monitored, and maintained at the target value, a 4σ quality level results in only 63 defects per million, and even a 3.5σ program results in just 465 defects per million parts. Table 4

Quality level	Mean Shift (in σ units)								
	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2
3.0σ	2700	3577	6440	12288	22832	40111	66803	105601	158700
3.5σ	465	666	1382	3011	6433	12201	22800	40100	66800
4.0σ	63	99	236	665	1350	3000	6200	12200	22800
4.5σ	6.8	12.8	32	88.5	233	577	1350	3000	6200
5.0σ	0.57	1.02	**3.4**	11	32	88.5	233	577	1300
5.5σ	0.034	.1056	.71	1.02	**3.4**	10.7	32	88.5	233
6.0σ	0.002	0.0063	0.019	.1	.39	1	**3.4**	11	32

Table 4: The number of defectives (ppm) for specified process mean shift and Quality levels

gives the number of nonconforming parts per million for different quality levels and different values of mean shift. If the process centering cannot be monitored or effectively controlled, a little latitude would be provided on each side of the specification to have some guard against process shifts.